



AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE
WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI,
INFORMATYKI I ELEKTRONIKI

ROZPRAWA DOKTORSKA

**Budowa optymalnych modeli uczenia na podstawie
wtórnych źródeł wiedzy**

Wiesław Paja

Promotor:
Prof. dr hab. inż. Zdzisław S. Hippe
Wyższa Szkoła Informatyki i Zarządzania
w Rzeszowie

KRAKÓW 2007

Pragnę wyrazić swoją wdzięczność promotorowi mojej pracy prof. Zdzisławowi S. Hippe, za opiekę naukową, wiele inspirujących dyskusji oraz konstruktywną krytykę.

Z całego serca dziękuję moim najbliższym za nieocenione wsparcie, którego mi udzielali.

*Niniejsza praca powstała, przy wsparciu finansowym Ministerstwa Nauki i Szkolnictwa Wyższego w ramach grantu promotorskiego **Nr 3 T11C 033 30** za co również serdecznie dziękuję.*

SPIS TREŚCI

| | |
|---|------------|
| Rozdział 1 | |
| Wstęp. Cel i zakres pracy | 4 |
| Rozdział 2 | |
| Reguły decyzji jako formalizm reprezentacji wiedzy | 9 |
| Rozdział 3 | |
| Indukcja reguł decyzji | 22 |
| Rozdział 4 | |
| Weryfikacja regułowej bazy wiedzy | 39 |
| Rozdział 5 | |
| Koncepcja optymalizacji wtórnych źródeł wiedzy | 51 |
| Rozdział 6 | |
| Generyczne operacje optymalizacyjne | 61 |
| Rozdział 7 | |
| Badania nad optymalizacją wtórnych źródeł wiedzy. Dyskusja wyników | 71 |
| Rozdział 8 | |
| Podsumowanie | 111 |
| Załącznik A | |
| Opis badanych baz informacyjnych | 113 |
| Literatura | 119 |

ROZDZIAŁ 1

WSTĘP. CEL I ZAKRES PRACY

Ogromna ilość informacji gromadzonych w postaci baz danych przez systemy informatyczne, stała się podstawą rozwoju jednego z nurtów sztucznej inteligencji zwanego *uczeniem maszynowym* (ang. *machine learning*) [Michalski, 1983]. Pojęcie to - najogólniej biorąc - dotyczy możliwości odkrywania i wyszukiwania informacji oraz wiedzy w danych, z zamysłem przedstawienia wyników tego procesu w postaci tzw. *modeli uczenia* (ang. *learning models*). Modele uczenia powstają zatem na podstawie danych historycznych (danych z przeszłości), i są następnie stosowane do objaśnienia danych nieznanymi (danych niewidzianych¹⁾). W relacji do jakości modeli uczenia, znany paradygmat sztucznej inteligencji mówi, że *zastosowanie różnych narzędzi nadzorowanego uczenia maszynowego do analizy danego (tego samego) zbioru danych, prowadzi zazwyczaj do utworzenia odmiennych modeli uczenia, objaśniających z różnymi skutkami i z różną precyzją ukrytą w nich wiedzę* [Winston, 1992]. Stosując tradycyjny sposób generowania modeli uczenia, powstają one wskutek działania algorytmów uczenia maszynowego, najczęściej nadzorowanego, zgodnie ze schematem indukcji logicznej. Modele te mają postać zależną od stosowanego algorytmu uczenia, i mogą być przedstawione w formie *symbolicznej* (jak np. formuły logiki predykatów pierwszego rzędu i rachunku zdań, reguły decyzji,

¹⁾ Pojęciem niewidzianych danych określa się takie dane, które nie brały udziału w procesie pozyskiwania wiedzy. Mogą to być nowe przypadki lub przypadki wydzielone z oryginalnych danych, służące do weryfikacji uzyskanych modeli.

ramki wiedzy, sieci semantyczne czy drzewa decyzji) lub w formie *niesymbolicznej* (m.in. sztuczne sieci neuronowe, algorytmy genetyczne) [Cichosz, 2000; Kubat, Bratko i Michalski, 1998]. Spośród tak szerokiego spektrum formalizmów reprezentacji wiedzy w dziedzinie uczenia maszynowego, metodą najbliższą do zapisu wiedzy stosowanego przez człowieka jest *regułowa reprezentacja wiedzy* w formie reguł JEŻELI ... TO. Modele uczenia tego typu są tematem szerszej dyskusji w [Cichosz, 2000; Michalski i inni, 1986; Mitchell, 1997; Stefanowski, 2001] oraz przedmiotem licznych badań. Regułowa reprezentacja wiedzy jest rezultatem wykonania procesu *indukcji reguł decyzji*. Proces ten jest jednym z podstawowych zadań uczenia maszynowego i systemów odkrywania wiedzy. Reguły decyzji, indukowane na podstawie zbiorów danych, reprezentują modele uczenia występujące w opisie *obiektów*, nazywanych również *przykładami uczącymi*. Opis zbioru przykładów (obiektów) może mieć postać *tablicy decyzji* wg Pawlaka [Pawlak, 1995], w której wiersze reprezentują kolejne przykłady, zaś kolumny odpowiadają wybranym cechom (atrybutom opisującym). Jeśli znana jest informacja o przynależności danego przypadku do określonej kategorii, zostaje to zapisane w ostatniej (skrajnie prawej) kolumnie tablicy decyzji.

Spośród metod bezpośredniej indukcji reguł decyzji, najbardziej znane są metody oparte na *sekwencyjnym pokrywaniu* (ang. *sequential covering*) zbioru przypadków uczących [Mitchell, 1997]. W rezultacie powstaje model uczenia w postaci zbioru reguł, pokrywających wszystkie przypadki ze zbioru uczącego. Obecnie znanych jest wiele algorytmów bezpośredniej indukcji reguł, m. in. algorytm *AQ* [Michalski i inni, 1986], *PRISM* [Cendrowska, 1987], *CN2* [Clark i Niblett, 1989], *LEM2* [Grzymała-Busse, 1997], *GTS (General-To-Specific)* [Hippe, 1997], *EACH* [Salzberg, 1991], *GOLEM* [Muggleton i De Raedt, 1994], czy algorytm *RISE (Rule Induction from a Set of Examples)* [Domingos, 1995; Domingos, 1996]. Istnieją także *pośrednie* metody generowania reguł decyzji, przekształcające pewne

formalizmy reprezentacji wiedzy w zbiór reguł decyzji. W tej grupie algorytmów najbardziej rozpowszechnione są metody konwersji drzew decyzji w zbiory reguł decyzji, np. algorytm ID3/C4.5 [Quinlan, 1993; Quinlan, 1996]. W metodzie generowania reguł decyzji z sieci przekonania [Hippe, Grzymała-Busse i Mroczek, 2004], sieć jest zamieniana na zbiór reguł przy uwzględnieniu wielkości wpływu poszczególnych atrybutów na decyzję. Szerzej znane są również metody ekstrakcji reguł przy użyciu algorytmów generujących sieci neuronowe [Duch, Adamczak i Grąbczewski, 1996; Duch, Adamczak i Grąbczewski, 1999]. W większości algorytmy indukcji reguł decyzji ukierunkowane są na klasyfikację nowych obiektów oraz generowanie minimalnych zbiorów reguł opisujących możliwie wszystkie przykłady uczące. Zazwyczaj główną wadą wspomnianych algorytmów jest duża liczba reguł w wygenerowanym modelu uczenia, co utrudnia przeciętnemu użytkownikowi percepcję oraz interpretację takiego modelu. Okazuje się, że bardzo duża liczba reguł w danym modelu może wynikać zarówno ze specyfiki stosowanych algorytmów uczenia maszynowego, jak i ze specyfiki analizowanej bazy przypadków.

Wspomniane wyżej cechy modeli uczenia w postaci zbioru reguł decyzji oraz przywołany uprzednio paradygmat sztucznej inteligencji stały się motywacją do podjęcia przeze mnie badań opisanych w tej dysertacji. Bliższa analiza dostępnych danych literaturowych wskazała bowiem, że różnorodność algorytmów, a także charakter danych mogą powodować powstawanie pewnych anomalii, wykrywanych w procesie zaawansowanej weryfikacji zbioru reguł. Zagadnieniem weryfikacji zbioru reguł zajmuje się wiele grup naukowych [Preece i Shinghal, 1994; Murrell i Plant, 1995; Spreeuwenberg i Gerrits, 2002; Ligęza, 2005], przy czym równocześnie (a także niezależnie) powstały różne systemy informatyczne do wykonywania tego typu procesów [Suwa, Scott i Shortliffe, 1982; Nguyen i inni, 1985; Ginsberg, 1988; Preece, 1989; Gerrits, Spreeuwenberg i Boekenoogen, 2000; Viktor, Paquet i le Roux, 2003].

Badania przedstawione w niniejszej pracy stanowią próbę opracowania metodologii optymalizacji wielu modeli uczenia, uzyskanych na podstawie tej samej tablicy decyzji. Zakres przeprowadzonych badań obejmuje dwa odrębne podejścia. W pierwszym z nich przewiduje się odrębną optymalizację modeli uczenia, utworzonych przy pomocy wybranych algorytmów uczenia maszynowego oraz porównanie ich cech i parametrów. Modele te nazwano dalej *wtórnymi źródłami wiedzy*. Natomiast w drugim podejściu, uzyskane modele uczenia łączone są w jedno globalne, wtórne źródło wiedzy, a następnie optymalizowane i porównywane z poszczególnymi modelami składowymi. W obydwu podejściach wszystkie modele uczenia generowane są na podstawie tej samej tablicy decyzji, nazywanej w pracy *pierwotnym źródłem wiedzy*, zaś zamysł optymalizacji modeli uczenia opiera się na koncepcji opracowania i zastosowania własnego algorytmu optymalizującego, wykonującego zestaw operacji generycznych na pierwotnym zbiorze reguł decyzji. Zakłada się, że operacje te doprowadzą do wyeliminowania zjawiska nadmiarowości w modelach uczenia, uzupełniając je nowymi, brakującymi regułami, a nawet spowodują utworzenie optymalnego (lub quasi-optymalnego) zbioru reguł decyzji. Do generowania reguł decyzji (tj. tworzenia wtórnych źródeł wiedzy) zastosowane będą trzy różne algorytmy uczenia maszynowego: *GTS* [Hippe, 1997], *LEM2* [Grzymała-Busse, 1997], *ID3/C4.5* [Quinlan, 1996]. Natomiast pierwotnymi źródłami wiedzy w prowadzonych badaniach będą niektóre bazy informacyjne zaczerpnięte z repozytorium baz informacyjnych UCI [Newman i inni, 1998], oraz własna baza informacyjna znamion melanocytowych skóry [Hippe i inni, 2003].

Wydaje się, że nakreślony tu zwięźle główny zamysł moich badań, metodyka ich realizacji oraz zakres, umożliwią zweryfikowanie zasadniczej tezy niniejszej rozprawy:

Genezą prowadzonych badań jest paradygmat sztucznej inteligencji, stwierdzający, że różne systemy uczenia maszynowego, zastosowane do analizy

tych samych danych, mogą tworzyć modele uczenia o odmiennych właściwościach, zwłaszcza w odniesieniu do wiarygodności oraz błędu klasyfikacji obiektów. Wydaje się zatem, że poprzez opracowanie odpowiednich algorytmów stapiania (łączenia) tych modeli, traktowanych jako wtórne źródła wiedzy, będzie można zdefiniować nową metodologię generowania modeli uczenia, cechujących się analogiczną, lub lepszą zdolnością klasyfikacji oraz identyfikacji obiektów, opisanych w logice atrybutalnej, w porównaniu do modeli generowanych w sposób tradycyjny.

ROZDZIAŁ 2

REGUŁY DECYZJI JAKO FORMALIZM REPREZENTACJI WIEDZY

Tematyką zagadnień poznania świata i wiedzy o nim zajmowali się od dawna filozofowie, a także logicy. Nowe spojrzenie na te problemy ujmuje *sztuczna inteligencja* (ang. *artificial intelligence*). W niniejszym rozdziale omówione zostanie pojęcie wiedzy oraz wybrane problemy związane z jej definicją, ujętą w kontekście metod sztucznej inteligencji. Na początku należy zdefiniować pojęcie *sztucznej inteligencji*, rozumianej często jako problem symulowania inteligentnych zachowań, czy też modelowania pracy mózgu. Najbardziej jednak adekwatną wydaje się być definicja zaproponowana przez Pawlaka [Pawlak, 1995], który stawia tej dziedzinie znacznie skromniejsze zadania. Mianowicie, głównym celem sztucznej inteligencji jest opracowanie algorytmów i metod, pozwalających na rozwiązanie możliwie szerokiej klasy rzeczywistych problemów, związanych z analizą danych pobieranych z otoczenia – i odpowiednie reagowanie na ich podstawie – celem rozwiązania określonego zadania. Tak więc inteligentne systemy mają na celu prawidłowe reagowanie na sygnały uzyskane z otoczenia. Jednak, żeby tego rodzaju system osiągał prawidłowe decyzje, koniecznym jest, by posiadał pewną wiedzę o otoczeniu w którym ma działać, lub innymi słowy, aby otoczenie było reprezentowane w jego pamięci w odpowiedni sposób. Zatem odpowiedni *formalizm reprezentacji wiedzy* w systemie nabiera szczególnego znaczenia. Okazuje się jednak, że problem do-

boru odpowiedniego formalizmu reprezentacji wiedzy jest o wiele starszy od technik informatycznych. Gromadzenie wiedzy było możliwe już od dawna dzięki wynalezieniu pisma. Reprezentacja wiedzy towarzyszy również rozwojowi techniki komputerowej od jej początku, gdzie przykładem może być odpowiedni zapis liczb w rejestrach komputera. Szybki rozwój tej dziedziny umożliwił usprawnienie istniejących oraz nowych metod reprezentacji wiedzy oraz przedstawienie starych zagadnień w nowym świetle. Formalizm reprezentacji wiedzy dotyczy więc sposobu przedstawienia wiedzy o otaczającym nas świecie oraz rodzaju wniosków, jakie możemy wyciągnąć na jej podstawie. Główne cechy wiedzy, które powinny być odzorowane w jej poprawnej reprezentacji [Kulikowski, 1995] to *wiarygodność* wiedzy, jej *historyczność* i *źródłowość*, *relatywność* oraz *strukturalność*.

Stopień *wiarygodności* wiedzy nie zawsze da się przedstawić w binarnej skali wartości logicznych. Czasem wygodniej posłużyć się skalą wielowartościową lub ciągłą, lub takimi nieostryimi określeniami jak „przypuszczalnie”, „na ogół”, „najczęściej”, itp. Wiedza ulega rozwojowi, tak co do zakresu, jak i co do trafności, dokładności opisu zjawisk, form opisu itp. W kontekście *historyczności* wiedzy istotnym jest rejestrowanie nie tylko jej głównych elementów, lecz także czas, w którym po raz ostatni były one weryfikowane. Istotną cechą wiedzy jest również *źródło*, z którego ją zaczerpnięto. W zależności od źródła bowiem, wiedza może być bardziej lub mniej wiarygodna. Kolejną, dość istotną właściwością wiedzy, jest jej *relatywność*. Nie zawsze bowiem pozornie prawdziwe stwierdzenia będą prawdziwe; można wykazać, że wiedza o danym problemie może być zupełnie inna, jeśli rozpatrzymy go w innych warunkach otoczenia, przyjmując np. inne założenia. Ostatnią wspomnianą cechą wiedzy jest jej *strukturalność*. Wiedza nie jest jedynie bezładnym nagromadzeniem stwierdzeń. Obszar wiedzy można dzielić tematycznie i to na wiele różnych sposobów. Występują w niej stwierdzenia ogólne, a także takie, które są ich szczegółowym

rozwinięciem. Strukturalizacja wiedzy ma coraz większe znaczenie w przypadku tworzenia systemów decyzyjnych dotyczących szerokiego zakresu tematycznego.

Najogólniej rzecz biorąc, opracowano różne formalizmy reprezentacji wiedzy, dostosowane do potrzeb systemów informacyjnych. Często stosowane jest tradycyjne rozróżnienie *symbolicznych* i *niesymbolicznych* (subsymbolicznych) metod reprezentacji wiedzy [Cichosz, 2000; Kubat, Bratko i Michalski, 1998]. Do pierwszej kategorii należą metody zorientowane na jawną reprezentację wiedzy (ang. *knowledge oriented methods*), wykorzystujące różnego rodzaju struktury przechowujące informację o charakterze symbolicznym, tj. zorganizowane w pewien sposób etykiety, którym można przypisać określoną interpretację. Taka wiedza jest na ogół zapisana w czytelnej dla człowieka postaci, i poddana może być dość łatwo jego kontroli. Zarówno całość, jak i poszczególne fragmenty symbolicznie reprezentowanej wiedzy, mają swoją interpretację w dziedzinie uczenia maszynowego. Najczęściej stosowanymi obecnie metodami symbolicznej reprezentacji wiedzy są formuły logiki predykatów pierwszego rzędu i rachunek zdań, reguły decyzji (ang. *decision rules*), ramki¹⁾ wiedzy (ang. *frames*), sieci semantyczne (ang. *semantic networks*), stwierdzenia, trójki <obiekt, atrybut, wartość> (ang. <*object, attribute, value*>) oraz drzewa decyzji (ang. *decision trees*).

W przypadku metod niesymbolicznych, zwanych również w literaturze *czarnymi skrzynkami* (ang. *black-box methods*), nie tworzących jawnej reprezentacji wiedzy, mamy do czynienia z informacją, której poszczególne elementy rozpatrywane osobno nie mają sensownej interpretacji. Są to najczęściej zbiory liczb lub łańcuchów binarnych, które łącznie reprezentują pewną wiedzę, lecz wiedza ta nie może być bezpośrednio wyrażona w po-

¹⁾ *Marvin Minsky* porównywał zaproponowany przez siebie formalizm reprezentacji wiedzy do klatek filmowych („... like in a movie ..”), czyli obiektów o małych rozmiarach [Minsky, 1974]. Z tego powodu lepiej w języku polskim oddaje sens tego pojęcia słowo „klatka wiedzy” lub „ramka wiedzy”, a nie jak się przyjęło w naszej literaturze informatycznej „rama wiedzy”.

staci zrozumiałej dla człowieka. Metody te odwołują się do obserwacji i doświadczeń zebranych na podstawie obserwacji otaczającego nas świata. Np. *sztuczne sieci neuronowe* symulują właściwości reprezentacji wiedzy i jej przetwarzania, podobnie jak komórki nerwowe zwierząt i ludzi. Sieć neuronowa to architektura obliczeniowa składająca się z prostych elementów przetwarzających – neuronów – które pracują równolegle i komunikują się ze sobą poprzez przesyłanie sygnałów [Krawiec i Stefanowski, 2003]. Inną technikę niesymbolicznej reprezentacji wiedzy stanowią *algorytmy genetyczne* [Rutkowska, Piliński i Rutkowski, 1997], które umożliwiają przekazywanie następnym generacjom wiedzy o całym gatunku; wiedza ta jest zapisana w genach. W kolejnych generacjach następuje poprawa cech całej populacji.

Wybór sposobu zapisywania wiedzy zależy od wielu czynników. Do najważniejszych należą: rodzaj wiedzy wymaganej do poprawnego działania systemu, niezbędna wielkość bazy wiedzy oraz unikanie redundancji informacji w bazie wiedzy. Najczęściej jednak stosowane są kombinowane sposoby zapisu wiedzy, polegające na jednoczesnym korzystaniu z różnych elementarnych form reprezentacji wiedzy. Sposób przedstawiania wiedzy w systemach informatycznych zwykle jest kompromisem pomiędzy potrzebą zapewnienia komunikatywności dla użytkownika, a koniecznością zagwarantowania efektywności przetwarzania wiedzy za pomocą środków technicznych i narzędzi programowych. W ocenie konstruowanych reprezentacji bierze się pod uwagę: (i) *wiarygodność* – czyli stopień, w jakim reprezentacja odpowiada rzeczywistości, (ii) *efektywność* – charakteryzuje przydatność reprezentacji do osiągnięcia danego celu, (iii) *poziom abstrakcji* – odpowiada zakresowi szczegółowości i precyzji pojęć używanych w reprezentacji; określając tzw. *moc opisową reprezentacji*.

Poszukiwana reprezentacja wiedzy powinna dodatkowo spełniać tzw. *postulat zrozumiałości* (ang. *comprehensibility*) [Michalski, 1983; Mitche,

Spiegelhalter i Taylor, 1994], tzn. powinna mieć formę, pozwalającą na zrozumienie, ocenę i dalsze użycie przez człowieka. Postulat ten jest uwzględniany szczególnie w przypadku metod symbolicznych [Eijkel, 1999].

Rodzaj użytego formalizmu reprezentacji wiedzy wraz z celem, któremu ta wiedza ma służyć, determinują sposób używania wiedzy. Do najbardziej typowych celów wykorzystania wiedzy należy zadanie *klasyfikacji* [Roiger i Geatz, 2003; Duda, Hart i Stork, 2001], tj. ustalania przynależności obiektów lub wzorców do określonej kategorii, a także *aproksymacja*, czyli odwzorowanie obiektów lub wzorców przy pomocy zbioru liczb rzeczywistych.

Niniejsza praca opiera się na regułowej reprezentacji wiedzy, dlatego też formalizm ten został szczegółowo zdefiniowany i omówiony w dalszej części rozdziału.

Regułowa reprezentacja wiedzy

Spośród szerokiego spektrum formalizmów reprezentacji wiedzy w dziedzinie uczenia maszynowego, metodą najbliższą do zapisu wiedzy stosowanego przez człowieka jest *regułowa reprezentacja wiedzy*. Zagadnienie to jest też tematem szerszej dyskusji w [Cichosz, 2000; Michalski i inni, 1986; Mitchell, 1997; Stefanowski, 2001] oraz treścią wielu badań, co pokrótce zostanie przedstawione w niniejszym rozdziale. Najpierw jednak należy scharakteryzować wspomniany formalizm i zdefiniować jego uogólnioną postać.

Reguły decyzji zwane są również regułami *składniowymi*, lub regułami *wnioskowania*. Istnieją dwa główne rodzaje reguł tego typu w zależności od ich źródła: *reguły fizyczne* (obiektywne) zdefiniowane dla człowieka przez naturę oraz *reguły logiczne* (subiektywne) zdefiniowane przez człowieka

w danej dziedzinie działania. Reguły fizyczne opisują naturalne zachowania i prawidłowości różnych systemów, które są znane z obserwacji i z doświadczeń. Mogą być dowodzone, są niezależne od woli człowieka, są uniwersalne i niezmiennie. Natomiast reguły logiczne, definiowane przez człowieka, są subiektywne; dotyczą systemów lokalnych, i w zależności od potrzeb są zmienne. Reguły fizyczne najczęściej opisują możliwe zachowania, mogą być stosowane w takich dziedzinach jak modelowanie, analiza i przewidywanie zachowań rozpatrywanych systemów. Natomiast reguły logiczne stanowią podstawę kształtowania zachowań człowieka, społeczności czy maszyn. W obu przypadkach definiowanie reguł logicznych musi uwzględniać również reguły fizyczne, których nie można pominąć; ponadto priorytet stosowania reguł jest wyższy dla reguł fizycznych.

Zagadnienie *indukcji reguł decyzji* dotyczy zbioru pojęć (zwanymi również *konceptami*, *kategoriami* czy też *klasami decyzyjnymi*) $K = \{K_j : j = 1, \dots, r\}$ w *tablicy decyzji* (ang. *decision table*) zawierającej niepusty i skończony zbiór przypadków uczących U . Należy zatem najpierw zdefiniować pojęcie *tablicy decyzji* [Pawlak, 1995]. Tablica decyzji zdefiniowana jest w postaci:

$$DT = (U, A \cup \{d\}) \quad (2.1)$$

gdzie A stanowi niepusty i skończony zbiór atrybutów, natomiast $d \notin A$ jest atrybutem decyzyjnym. Zbiór V_a jest dziedziną atrybutu $a \in A$, $V = \bigcup_{a \in A} V_a$ oraz $f: U \times A \rightarrow V$ jest funkcją informacyjną, taką że

$\forall_{a \in A, x \in U} f(a, x) \in V_a$. Jeżeli tablica decyzji zawiera przykłady pojęć, to dla każdego przykładu $x \in U$ wartość funkcji informacyjnej $f(d, x) \in K$.

Dla każdego pojęcia K_j podaje się zbiór przykładów uczących E_{K_j} złożony z obiektów $x \in U$ opisanych ciągiem wartości atrybutów warunkowych $a_i \in A$, tj. $(f(a_1, x), f(a_2, x), \dots, f(a_m, x))$ oraz wartością atrybutu decyzyjnego $f(d, x) = K_j$.

W procesie indukcji reguł decyzji zbiór obiektów U dzielony jest na niepuste podzbiory E_{K_i}, E_{K_j} parami rozłączne, tzn. $\forall i \neq j E_{K_i} \cap E_{K_j} = \emptyset$ oraz $\bigcup_{j=1}^r E_{K_j} = U$. Niech $U = E_{K_j}^+ \cup E_{K_j}^-$ dla danej kategorii K_j oznacza podział zbioru U na podzbiór przykładów pozytywnych $E_{K_j}^+ = E_{K_j}$ oraz przykładów negatywnych $E_{K_j}^- = U \setminus E_{K_j}$ (tj. pozostałych przykładów).

Ogólna postać regułowej reprezentacji wiedzy składa się z części warunkowej P oraz części decyzyjnej Q . Część warunkowa jest zbiorem (*kompleksem*) tzw. *warunków elementarnych*, będących kombinacją atrybutu opisującego i jego wartości. Zatem reguła r opisująca klasę decyzyjną K_j , ma formę wyrażenia

$$\mathbf{JEŻELI } P \text{ TO } Q \quad (2.2)$$

Zastosowanie reguł takiej postaci w dziedzinie sztucznej inteligencji oraz inżynierii wiedzy (ang. *knowledge engineering*) zostało zaczerpnięte z dziedziny logiki matematycznej [Ligęza, 2005]. Zbiór wygenerowanych reguł stanowi bowiem pewien zbiór logicznych implikacji typu

$$\mathbf{P} \rightarrow \mathbf{Q} \text{ lub odpowiednio } p_1 \wedge p_2 \wedge \dots \wedge p_n \rightarrow Q \quad (2.3)$$

gdzie część warunkowa P jest koniunkcją warunków elementarnych p . Liczbę wykorzystywanych warunków nazywa się często *długością reguły* [Stefanowski, 2001]. Warunek elementarny może mieć postać następującą:

atrybut JEST wartość_atrybutu

Słowo *JEST* pełni rolę słowa kluczowego, oddzielającego nazwę atrybutu od jego wartości i określającego ich wzajemną relację. Można wyróżnić również inne operatory relacji, jak np. SA, =, ≠, <, >, ≤, ≥. Całą postać reguły można zatem przedstawić następująco:

JEŻELI warunek 1

ORAZ warunek 2

...

ORAZ warunek n

TO decyzja

gdzie słowa *JEŻELI*, *ORAZ* i *TO* pełnią rolę słów kluczowych. Słowo *JEŻELI* (zamiennie *JEŚLI*, *GDY*, *KIEDY*, itp.) rozpoczyna regułę, słowo *ORAZ* (zamiennie *I*, *TAKŻE*, itp.) rozdziela poszczególne warunki w części warunkowej reguły, natomiast słowo *TO* (zamiennie *WTEDY*, itp.) oddziela część warunkową reguły od części decyzyjnej.

W niektórych przypadkach używa się reguł w formie **M-z-N** [Towell i Shavlik, 1993] tzn. takich, których przesłanki zawierają zdania typu „*M spośród N podanych warunków jest spełnionych*”. Tego typu sformułowania, podobnie jak miary odległości, pozwalają wyrażać koncepcje dotyczące jednocześnie wielu warunków (na przykład w opisie wyników głosowania można w zwarty sposób wyrazić stwierdzenie „większość jest za”), które trudno wyrazić prostymi regułami logiki klasycznej.

Reguła decyzji pozwala określić decyzję przy spełnieniu określonych warunków w danej sytuacji. Prostota interpretacji, a jednocześnie duża siła wyrazu reguł, powoduje ich ogromną popularność w zastosowaniach z użyciem formalizmów reprezentacji wiedzy. W niektórych zastosowaniach reguły służą również do odwzorowania i przedstawienia innych, bardziej skomplikowanych struktur reprezentacji wiedzy jak drzewa decyzji, sieci przekonań, sieci neuronowe [Quinlan, 1993; Hippe, Grzymała-Busse i Mroczek, 2004; Duch, Adamczak i Grąbczewski, 1996; Śnieżyński, 2006; Setiono, 1996].

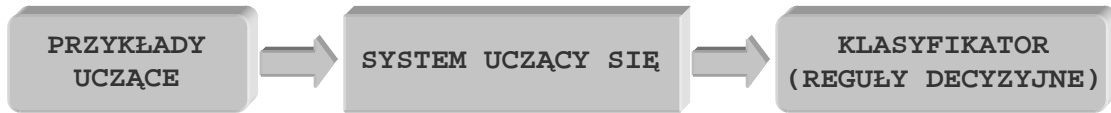
Zgodnie z poglądami przedstawionymi w [Żytkow, 2002] reguły wygenerowane z danych można podzielić na trzy rodzaje: *reguły klasyfikacyjne*, *reguły charakterystyczne* oraz *reguły asocjacyjne*. Pierwsze z nich związa-

ne są z klasycznym, indukcyjnym procesem uczenia, i określają czy przypadek spełniający warunki elementarne, znajdujące się w *części warunkowej* reguły, należy do klasy decyzyjnej zawartej w *części decyzyjnej* danej reguły. Reguły charakterystyczne (ang. *characteristic rules*) wskazują najbardziej charakterystyczne własności obiektów należących do danej klasy obiektów [Turmeaux i inni, 2003; Brijs, Vanhoof i Wets, 2000]. Na przykład, aby przedstawić fakt, że łabędź jest ptakiem i ma zazwyczaj biały kolor można zastosować następującą regułę charakterystyczną: *JEŻELI gatunek = łabędź TO rodzaj = ptak ORAZ kolor = biały*. Natomiast reguły asocjacyjne (ang. *association rules*) opisują wiedzę o zależnościach wystąpienia wartości niektórych atrybutów łącznie z wartościami innych atrybutów [Zygmunt, 2001], np. „58% klientów kupujących kosmetyki, kupuje również czasopisma”. Kulikowski [Kulikowski, 2006] proponuje teoretyczny model wnioskowania opartego na tzw. *regułach deontycznych* mających postać: *Jeśli A, to uczeń B, aby C*, przy czym *A* i *C* są zdaniami orzekającymi, określającymi odpowiednio istniejący i pożądany stan w pewnym obszarze świata zewnętrznego, a *B* jest regułą działania, które stanowi opisanemu przez *A*, nadaje cechy scharakteryzowane przez *C*. W niniejszej pracy głównym rozpatrywanym zagadnieniem w kontekście rodzaju reguł są reguły klasyfikacyjne.

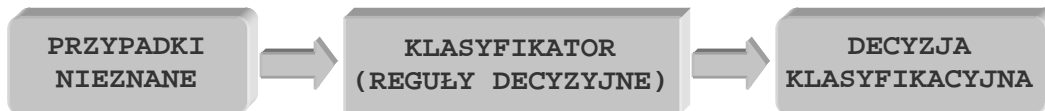
Regułowa reprezentacja wiedzy jako mechanizm działający w systemach ekspertowych

W niniejszej pracy rozpatrywany jest kontekst *nadzorowanego* uczenia maszynowego (ang. *supervised learning*), tzn. takiego w którym system uczący się otrzymuje zbiór przykładów w postaci wektorów zawierających dwa elementy: informacje wejściowe do systemu i odpowiadające im pożądane informacje wyjściowe (odpowiedzi systemu). Schemat obrazujący zastosowanie reguł decyzji w takim systemie przedstawiono na Rys. 2-1 oraz

Rys. 2-2. Zbiór reguł decyzji wyznaczany jest na podstawie informacji dostarczanych przez skończony zbiór treningowy składający się z obiektów o znanych wartościach atrybutów.



Rys. 2-1 Uczenie klasyfikatora na podstawie zbioru przykładów uczących



Rys. 2-2 Klasyfikowanie nowych, nieznanymi przypadków

System uczący się poszukuje klasyfikatora zwanego również modelem uczenia (ang. *learning model*), którym w tym przypadku jest zbiór reguł. Utworzony klasyfikator przy zastosowaniu wybranych metod klasyfikacji, przyporządkowuje skończony zbiór przypadków nieznanymi do określonych kategorii decyzyjnych. Spośród systemów stosujących wiedzę w postaci reguł lub w innej, równoważnej formie, można wymienić:

- systemy generujące reguły metodą znajdowania pokryć: **AQ** [Michalski, 1969], **CN2** [Clark i Niblett, 1989; Clark i Boswell, 1991], **GTS** (General-To-Specific) [Hippe, 1997], **PRISM** [Cendrowska, 1987],
- systemy których podstawą konstrukcji jest teoria zbiorów przybliżonych [Pawlak, 1991]: **LEERS** (Learning from Examples Based on Rough Sets) [Grzymała-Busse, 1992], **Rosetta** [Øhrn i inni, 1998],
- systemy posługujące się pojęciem drzewa decyzji: **CART** (Classification And Regression Trees) [Breiman i inni, 1984], **ID3**, **C4.5**, [Quinlan, 1993; Quinlan, 1996], **C5.0** [Quinlan, 1997], **GIMS** [Galant, 1996; Bonner, Galant i Tyburcy, 2001], **Quest** [Loh i Shih, 1997],
- systemy generujące reguły z użyciem sieci przekonań np. **BeliefSEEKER** [Hippe, Grzymała-Busse i Mroczek, 2004; Śnieżyński, 2006], z użyciem

sieci neuronowych [Duch, Adamczak i Grąbczewski, 1996] algorytm **SUBSET** [Towell i Shavlik, 1993] czy **NeuroRule** [Setiono, 1996], przy użyciu *algorytmów genetycznych* [Carvalho i Freitas, 2004; Kwedlo i Krętowski, 2002], czy też przy użyciu *analizy skupień* [Haendel, 2003].

Interpretacja reguł

Zadanie interpretacji i aktywowania (ang. *firing rules*) reguł należy do *mechanizmu wnioskującego*. Użycie reguły składa się z następujących kroków: (i) sprawdzenie czy spełnione są warunki zawarte w regule, (ii) podjęcie decyzji (lub podjęcie odpowiedniego działania), (iii) powrót do mechanizmu wnioskującego.

Proces zastosowania zbioru reguł można podzielić jak niżej: (i) selekcja dostępnych reguł, (ii) wykrycie zbioru reguł możliwych do uaktywnienia (ang. *triggered rules*), tzw. *conflict set*, (iii) wybór konkretnej reguły do klasyfikacji.

Proces wnioskowania jest przeprowadzany dwoma metodami: *wnioskowanie wprzód* (ang. *forward chaining reasoning*) i *wnioskowanie wstecz* (ang. *backward chaining reasoning*). Może być również stosowane wnioskowanie mieszane. Najpopularniejsza strategia wnioskowania polega na użyciu wnioskowania wprzód jako głównej strategii, a na pewnych poziomach, jeżeli poszukiwany jest szczegółowy cel, jest stosowane wnioskowanie wstecz [Nalepa, 2004; Ligęza, 2005].

Strategia wnioskowanie wprzód jest również znana pod nazwą metody sterowania danymi (ang. *data-driven*). W tej strategii reguły są stosowane do faktów i do reguł, zgodnie z odkrywaną nową wiedzą. Algorytm kończy działanie, gdy cel jest osiągnięty. Wnioskowanie wprzód może być stosowane praktycznie w każdym przypadku. Jest szczególnie użyteczne wtedy, gdy istnieje relatywnie niewiele faktów początkowych, lub nowe fakty mu-

szą być utworzone, lub też gdy cel jest luźno określony lub nie jest określony w całości. Zwykle wnioskowanie wprzód pozwala utworzyć dobrze zdefiniowane fakty podstawowe.

Strategia wnioskowania wstecz, zwana również metodą *sterowaną przez cel* (ang. *goal-driven*), wnioskowaniem *zstępującym* lub też *regresywnym*, polega na wnioskowaniu w kierunku przeciwnym do kierunku wynikania logicznego określonego zbioru reguł decyzji. Rozpatrywane stwierdzenie jest konkluzją reguły, i dla wykazania jego prawidłowości, badane są przesłanki mogące być z kolei konkluzjami innych reguł. Podczas realizacji tego typu wnioskowania, podejmowane są próby uaktywniania tych reguł, których konkluzje są przesłankami reguły prowadzącej do badanej hipotezy. W tym wnioskowaniu prawa strona reguły (konkluzja) jest punktem startowym i przesłanki wyszukuje się po lewej stronie reguły.

W procesie wnioskowania można spotkać tzw. *rozwiązywanie konfliktów* (ang. *conflict resolution*), które występuje w przypadku, kiedy można aktywować więcej niż jedną regułę umożliwiającą osiągnięcie celu wnioskowania. Problem ten nie istnieje w systemach deterministycznych, tzn. takich, w których tylko jedna reguła może być zastosowana w jakimkolwiek etapie. W większości systemów wybierana jest jedna reguła według pomocniczych kryteriów, innych niż kryteria służące do spełnienia warunków w części warunkowej. Podstawowe określenie rozwiązania dotyczy regułowych systemów kontrolnych składających się z n reguł. Przyjęto, że reguły są uporządkowane liniowo zgodnie z ich wskazaniem. Rozwiązanie konfliktu polega wówczas na wybraniu dokładnie jednej reguły, stosowanej w danej kategorii. Wybrana reguła zostaje zastosowana, i proces zaczyna się od początku. Podstawowe strategie rozwiązania powyższego problemu mogą zależeć od specyfikacji i zastosowania systemu. Istnieją dwie podstawowe możliwości: (i) jako pierwsza może być wybrana reguła odpowiednia do osiągnięcia pożądanego w danym momencie celu, a następnie

może być zapewnione spełnienie warunków tejże reguły, (ii) reguły są sekwencyjnie testowane czy spełniane są ich przesłanki, i pierwsza spełniona reguła jest wybierana i wykonywana. Rozwiązanie pierwsze jest wykorzystywane w większości znanych systemów. Wymienia się również inne znane strategie: (i) uporządkowanie wg reguł (reguły pojawiające się najwcześniej posiadają najwyższy priorytet), (ii) uporządkowanie wg danych (reguła z najwyższym priorytetem przypisanym do jej warunków posiada najwyższy priorytet), (iii) uporządkowanie wg rozmiaru (reguła z najdłuższą listą spełnionych warunków ma najwyższy priorytet), (iv) uporządkowanie wg specyficzności (szeregowanie reguł których warunki zawierają inne reguły), (v) ograniczenie kontekstowe (polega na aktywacji lub deaktywacji grup reguł w każdej chwili aby zredukować wystąpienie konfliktu);

Powyższa charakterystyka reguł decyzji nakreśliła ich rolę jako szeroko stosowanego w uczeniu maszynowym formalizmu reprezentacji wiedzy. W kolejnym natomiast rozdziale skupiono się na procesie tworzenia reguł decyzji, parametrach charakteryzujących reguły oraz na procesie klasyfikacji przypadków nieznanych przy pomocy zbiorów reguł decyzji.

ROZDZIAŁ 3

INDUKCJA REGUŁ DECYZJI

Reguły decyzji najczęściej są generowane *bezpośrednio* z bazy przypadków lub *pośrednio*, wykorzystując do tego celu inne formalizmy reprezentacji wiedzy. Spośród znanych metod bezpośredniej indukcji reguł decyzji, najbardziej znane są metody oparte na *sekwencyjnym pokrywaniu* (ang. *sequential covering*) zbioru przypadków uczących, natomiast mniej znane są metody korzystające ze strategii *pokrywania równoległego* (ang. *simultaneous covering*). W metodzie pokrywania sekwencyjnego system generuje najpierw pierwszą regułę, następnie w następnym cyklu iteracji pomija przypadki *pozytywne* (przypadki, które spełniają zarówno część warunkową, jak i część decyzyjną reguły), które ta reguła pokrywa. W następnym kroku proces ten zostaje powtórzony dla pozostałych przypadków w analizowanej bazie informacyjnej. W rezultacie powstaje zbiór reguł, które pokrywają wszystkie przypadki ze zbioru uczącego. Ogólna postać omawianego algorytmu [Mitchell, 1997] przedstawiona została poniżej:

```
Sekwencyjne pokrywanie ( $K_j$  - klasa;  $A$  - atrybuty;  $E$  - przykłady;  $\tau$  - próg akceptacji);  
 $R := \emptyset$ ; {zbiór poszukiwanych reguł}  
 $r := \text{learn-one-rule}(\text{klasa } K_j; \text{ atrybuty } A; \text{ przykłady } E)$   
while  $\text{evaluate}(r, E) > \tau$  do  
   $R := R \cup r$ ; {dodaj znalezioną regułę  $r$  do zbioru poszukiwanych reguł  $R$ }  
   $E := E \setminus [R]$ ; {usuń przykłady pozytywne, pokryte przez  $R$ }  
   $r := \text{learn-one-rule}(\text{klasa } K_j; \text{ atrybuty } A; \text{ przykłady } E)$   
return  $R$ 
```

Najistotniejszym elementem powyższego algorytmu jest funkcja *learn-one-rule*, tworząca regułę, która prawidłowo pokrywa możliwie największą liczbę przykładów, a jednocześnie pokrywa nieprawidłowo jak najmniejszą liczbę przykładów. Typowe opisy działania omawianego algorytmu można znaleźć w [Cendrowska, 1987; Clark i Boswell, 1991; Grzymała-Busse, 1992; Michalski, 1983; Hippe, 1997]. Większość z tych algorytmów działa w oparciu o podejście zwane *general-to-specific (top-down method)*, polegające na wyborze początkowego warunku dla tworzonej reguły, a następnie dalszej jej specjalizacji, poprzez dodawanie kolejnych warunków elementarnych. Specjalizacja reguły trwa do momentu osiągnięcia kryterium akceptacji reguły (funkcja ewaluacyjna, ang. *evaluate*). Kryterium tym może być np. *niesprzeczność*, co znalazło implementację w algorytmach *AQ* [Michalski i inni, 1986], *GTS* [Hippe, 1997], *LEM2* [Grzymała-Busse, 1997] oraz *PRISM* [Cendrowska, 1987]. Z kolei funkcja *evaluate* jest zależnym od użytkownika kryterium pozwalającym oszacować jakość tworzonego zbioru reguł a tym samym określa zakończenie procesu generowania reguł, np. kolejne reguły tworzone są nieprzerwanie, dopóki w zbiorze uczącym istnieją przykłady pozytywne z klasy K_j , nie pokryte przez żadną z wygenerowanych reguł [algorytmy *AQ*, *GTS*, *LEM2*, oraz *PRISM*]. Natomiast inne algorytmy umożliwiają określenie warunku wcześniejszego zakończenia poszukiwania zbioru reguł poprzez ustawienie odpowiedniej wartości progu τ [Clark i Boswell, 1991]. W takim przypadku algorytm generuje reguły do momentu osiągnięcia założonej wartości parametru progowego, np. zakładany poziom procentowy pozytywnie pokrywanych przypadków (ang. *minimum purity*) [Fürnkranz, 1999] czy też poziom różnicy między liczbą przypadków pokrywanych przez regułę a liczbą przypadków w zbiorze uczącym (ang. *significance test*).

Istnieją również algorytmy działające w oparciu o metodę przeciwną do omówionej powyżej, a mianowicie metodę *specific-to-general (bottom-up)*

method). Przykładem takiego algorytmu jest algorytm *RISE* (**R**ule **I**nduc**T**ion from a Set of **E**xamples) [Domingos, 1995; Domingos, 1996], *EACH* [Salzberg, 1991] czy też algorytm *GOLEM* [Muggleton i De Raedt, 1994]. Zaletą tych algorytmów jest tworzenie reguł będących dokładniejszym opisem rozpatrywanego problemu, niż opis osiągnięty przy pomocy reguł generowanych według schematu *general-to-specific*. W przypadku gdy cel badań (poszukiwana konkluzja) formalnie znajduje się daleko od punktu startowego (wybranego atrybutu startowego), od którego zaczyna się budowa reguły, jego znalezienie może okazać się trudne. Rośnie ryzyko zbudowania złej ścieżki do celu, i rośnie prawdopodobieństwo uzyskania nieprawidłowych rezultatów. Jeżeli część opisu danego konceptu jest wystarczająco szczegółowa, tzn. jeśli zawiera właściwą liczbę atrybutów opisujących przypadki, algorytmy takie jak *C4.5* czy *CN2* [Clark i Niblett, 1989] mogą ich nie znaleźć, lub mogą zatrzymać przeszukiwanie zbyt wcześnie. Metoda ta posiada dodatkowe wady: poszukiwanie musi rozpocząć się od szczegółowego przypadku, powodując dużą wrażliwość na zakłócenia, co w efekcie prowadzi do tego, że końcowe decyzje mogą być nieprawidłowe i trudne do korekty. W literaturze można również znaleźć informacje na temat algorytmów łączących obydwa podejścia [Domingos, 1995], jednakże zarówno te algorytmy jak i algorytmy *specific-to-general* nie znalazły powszechnego zastosowania w dziedzinie uczenia maszynowego.

Algorytm *sekwencyjnego pokrywania* może być stosowany w odniesieniu do całego zbioru przypadków uczących (*CN2*, *GTS*) lub oddzielnie dla każdej klasy K_j (*AQ*, *LEM2*, *PRISM*). W przypadku rozpatrywania całości zbioru należy zwrócić uwagę na odpowiednie rozróżnienie przykładów pozytywnych i negatywnych. Wspomniany algorytm reprezentuje heurystykę zachłanną, polegającą na sukcesywnym pomijaniu w kolejnych cyklach iteracyjnych przykładów uczących pokrywanych przez tworzone reguły. W podobny sposób traktowane są również warunki elementarne wybierane do budowy reguły.

Jednym z najważniejszych elementów w procesie budowy reguł decyzji jest ocena kandydatów na części warunkowe reguł. Jest to jedna z cech różniących popularne algorytmy generujące reguły. Szczegółowo zagadnienie to przedstawione jest w następnym podrozdziale.

Kryteria oceny jakości reguł

Systemy indukcji reguł generują reguły decyzji na podstawie zbiorów uczących. Zbiór utworzonych reguł określa działanie klasyfikatora wykorzystującego te reguły do klasyfikacji obiektów nieznanymi. Istotnym jest aby generowane reguły decyzji posiadały wysoką zdolność predykcji oraz odpowiednie dopasowanie do rozpatrywanego problemu. Właściwości te są określane przy pomocy funkcji zwanej *jakością reguł*. Pomiar jakości reguł jest niezbędny w procesie indukcji reguł oraz w procesie klasyfikacji. Proces indukcji reguł jest zwykle rozumiany jako przeszukiwanie przestrzeni hipotez możliwych reguł, w celu znalezienia reguły decyzji spełniającej określone kryteria. W procesie indukcji reguł opartym na metodzie poszukiwania *general-to-specific* (CN2 [Clark i Boswell, 1991], GTS [Hippe, 1997], HYDRA [Ali i Pazzani, 1993]) jakość reguł może być stosowana do oceny wyboru par *atrybut-wartość* w procesie budowy reguł i/lub może być również ważną miarą wstrzymującą dalszą specjalizację budowanej reguły. Wstrzymanie budowy reguły związane jest z tzw. *problemem małej dysjunkcji* (ang. *small disjunct problems*), szerzej badanym i opisywanym w [An i Cercone, 2001; Weiss i Hirsh, 2000].

Miara jakości reguł jest również niezbędna w procesie klasyfikacji obiektów. Zdarza się bowiem, że przypadek nieznanymi pokrywany jest przez wiele reguł, wskazujących jego przynależność do odmiennych kategorii. W tej sytuacji, należy zastosować odpowiedni schemat postępowania w celu przypisania przypadku do właściwej kategorii. Zrealizowanie tego celu mo-

że wymagać zdefiniowania numerycznego parametru, określającego moc klasyfikacyjną reguły, jej wiarygodność, itp.

Miary oceny jakości reguł

Wiele miar oceny jakości reguł wywodzi się z analizy relacji pomiędzy regułą decyzji R a klasą decyzyjną C. Relacje te mogą być przedstawiane w postaci tzw. *tablicy kontyngencji* (ang. *contingency table*) o wymiarze 2 x 2 [Bruha i Kockova, 1993], która zawiera wartości jak w Tabelicy 3-1, gdzie N_{rc} oznacza liczbę przypadków uczących *pokrywanych* przez regułę R i *należących* do klasy C, $N_{r\bar{c}}$ oznacza liczbę przypadków uczących *pokrywanych* przez regułę R ale *nie należących* do klasy C, N_r oznacza *całkowitą* liczbę przypadków uczących, $N_{\bar{r}c}$ oznacza liczbę przypadków uczących *nie pokrywanych* przez regułę R ale *należących* do klasy C, $N_{\bar{r}\bar{c}}$ oznacza liczbę przypadków uczących *nie pokrywanych* przez regułę R i *nie należących* do klasy C, $N_r, N_{\bar{r}}, N_c, N_{\bar{c}}$ oznaczają sumy marginalne, np. $N_r = N_{rc} + N_{r\bar{c}}$, czyli liczba przypadków *pokrywanych* przez regułę R.

Tablica 3-1 Tablica kontyngencji

| | Klasa C | Klasa inna od C | |
|------------------------------|----------------|----------------------|---------------|
| Pokrywane przez regułę R | N_{rc} | $N_{r\bar{c}}$ | N_r |
| Nie pokrywane przez regułę R | $N_{\bar{r}c}$ | $N_{\bar{r}\bar{c}}$ | $N_{\bar{r}}$ |
| | N_c | $N_{\bar{c}}$ | N |

Tablica kontyngencji może być również przedstawiona przy użyciu względnych wartości częstości wystąpień, co przedstawiono w Tablica 3-2, gdzie

$$f_{rc} = \frac{N_{rc}}{N}; \quad f_{r\bar{c}} = \frac{N_{r\bar{c}}}{N}; \quad \text{itd.} \quad (3.1)$$

Tablica 3-2 Tablica kontyngencji zawierająca względne wartości częstości wystąpień

| | Klasa C | Klasa inna od C | |
|------------------------------|----------|-----------------|-------|
| Pokrywane przez regułę R | f_{rc} | f_{rc} | f_r |
| Nie pokrywane przez regułę R | f_{rc} | f_{rc} | f_r |
| | f_c | f_c | 1 |

Zgodnie z podziałem przedstawionym w [An i Cercone, 2001], miary oceny jakości reguł można podzielić na kilka kategorii, a mianowicie na:

- *formuły empiryczne* - które stanowią podejście tzw. *ad hoc* do definicji jakości reguły. Bruha [Bruha i Kockova, 1993] określa takie formuły jako empiryczne, ponieważ nie mają one podstaw teoretycznych. Są one określane na podstawie intuicji badawczej. Przykładem takich miar są np.: *jakość opisu* (ang. *description quality*) [Michalski i Kaufman, 1999], *wagowa suma konsystencji i pokrycia* (ang. *weighted sum of consistency and coverage*) [Michalski, 1990], *iloczyn konsystencji i pokrycia* (ang. *product of consistency and coverage*) [Brazdil i Torgo, 1990], *miara H* będąca kombinacją miary *ogólności* reguły oraz jej *dokładności* [Hippe, 1997],

- *miary asocjacji* – które dotyczą relacji pomiędzy klasyfikacją dla kolumn a klasyfikacją dla wierszy w tablicy kontyngencji. Można wyróżnić tutaj kilka statystyk używanych do pomiaru asocjacji np.: *statystyka Pearsona* χ^2 [Bruha i Kockova, 1993], *statystyczny stosunek prawdopodobieństw* G^2 (ang. *likelihood ratio statistic*) [Clark i Niblett, 1989], *m-estymata prawdopodobieństwa* [Cestnik, 1990], *estymata Laplace'a* [Clark i Boswell, 1991].

- *miary zgodności* (ang. *measure of agreement*) - dotyczące asocjacji elementów w tablicy kontyngencji tylko na jej przekątnej diagonalnej. Można tu przytoczyć przykłady dwóch miar zgodności przedstawione w literaturze: *formuła Cohena* [Cohen, 1960], *formuła Colemana* [Bishop, Fienberg i Holand, 1991; Bruha i Kockova, 1993].

- *miary informacji* – definiowane przede wszystkim przez *entropię informacji*: $Ent(S) = -\sum_{i=1}^r p_i \cdot \log_2 p_i$, gdzie p_i jest prawdopodobieństwem wystąpienia klasy K_i w zbiorze przykładów S pokrytych przez koniunkcję warunków P , natomiast r jest liczbą klas decyzyjnych [Cendrowska, 1987; Clark i Boswell, 1991], *information score* - statystyczna miara używana do określenia jakości reguł. Miara ta określa ilość informacji jaką wnosi do zagadnienia dana reguła R [Kononenko i Bratko, 1991]. Często stosowana jest także miara *J-measure* – bazująca na prawdopodobieństwie danej klasy oraz entropii informacji wewnątrz danej klasy decyzyjnej. Stosowana jest w algorytmie *ITrule* do oceny kandydatów na części warunkowe reguł decyzji [Feng i Michie, 1994].

- *miary logicznej wystarczalności* (ang. *logical sufficiency*) – która jest standardowym, statystycznym stosunkiem prawdopodobieństw, który został zastosowany do pomiaru jakości reguły [Duda, Gaschnig i Hart, 1979; Ali i Pazzani, 1993].

- *miara dyskryminacji* (ang. *discrimination*) – będąca kolejną statystyczną miarą oceny reguł. Miara ta określa zakres do którego warunki zapytania odróżniają istotne i nieistotne dokumenty [Robertson i Sparck Jones, 1976].

Michalski do oceny kandydatów na części warunkowe reguły proponuje rozważyć szereg kryteriów [Michalski i inni, 1986], takich jak maksymalizacja liczby przykładów pozytywnych pokrywanych przez koniunkcję P , maksymalizacja ilorazu liczby przykładów pozytywnych pokrywanych przez koniunkcję P przez ogólną liczbę wszystkich przykładów, a także minimalizacja liczby użytych warunków elementarnych. Np. algorytm *AQ* [Michalski, 1969] wykorzystuje do budowy reguł decyzji tzw. *gwiazdę* specjalizującą maksymalnie ogólne warunki złożone, pokrywające wybrany początkowy przykład pozytywny *SEED*. Algorytm upewnia się że żaden z negatywnych przypadków nie jest pokrywany przez budowaną regułę. Następnie usuwa wszystkie pozytywne przypadki które mogą być pokryte

przez regułę. Działanie jest kontynuowane z pozostałymi przypadkami pozytywnymi. Algorytm *AQ* przedstawiono poniżej.

```

P - zbiór przykładów pozytywnych klasy K
N - zbiór przykładów negatywnych klasy K
Procedure AQ(P, N)
    R := ∅;                                {zbiór poszukiwanych reguł}
    while P ⊄ [R] do                    {zbiór reguł R nie pokrywa wszystkich przypadków zbioru P}
        SEED := p ∈ P;                       {SEED - pozytywny przypadek p nie pokrywany przez R}
        STAR := STAR(SEED | N) {zbiór kompleksów pokrywających SEED ale nie pokrywających przypadków ze zbioru N}
        BEST := b ∈ STAR                     {wybierz b najlepszy element ze zbioru STAR, b jest najkrótsza, najprostszą regułą}

        R := R ∪ BEST
    return R.

Procedure STAR(SEED, N)
    STAR - zbiór zawierający pusty kompleks
    while istnieje kompleks w zbiorze STAR pokrywający przypadki negatywne ze zbioru N,
        Wybierz przypadek negatywny Eneg pokrywany przez kompleks ze zbioru STAR.
        Specjalizuj kompleksy w zbiorze STAR aby wyłączyć Eneg:
            EXTENSION - wszystkie selektory pokrywające SEED ale nie Eneg
            Niech STAR będzie zbiorem {x∧y / x ∈ STAR, y ∈ EXTENSION}.
            Usuń wszystkie kompleksy w STAR pochłaniane przez inne kompleksy
        repeat until rozmiar STAR ≤ maxstar (maximum określone przez użytkownika).
            Usuń najgorszy kompleks ze zbioru STAR.
    return STAR.

```

Algorytmy *PRISM* i *CN2* do oceny koniunkcji warunków *P* stosują wspomnianą powyżej miarę *entropii informacji*, choć pewne wersje *CN2* wykorzystują *m-estymatę*. Algorytm *CN2* łączy w sobie podejścia znane z algorytmu *AQ* oraz z metodologii budowy drzew decyzji. Algorytm ten wykorzystywany jest w wielu systemach odkrywania wiedzy np. w *MLC++*

[Kohavi, Sommerfield i Dougherty, 1996], *Weka* [Garner, 1995]. Pseudokod algorytmu *CN2* przedstawiony został poniżej.

```

E - zbiór przykładów uczących
Procedure CN2(E)
    RULE_LIST :=  $\emptyset$ ; {zbiór poszukiwanych reguł}
    repeat
        BEST_CPX := Find_Best_Complex(E);
        if BEST_CPX  $\neq$  0 then
            Niech E' będzie zbiorem przykładów pokrytych przez
            BEST_CPX;
            E := E \ [E']; {usuń ze zbioru E przykłady E' pokryte
                przez BEST_CPX}
            Niech C będzie najczęstsza klasą w przykładach E';
            Niech r będzie regułą postaci 'if BEST_CPX then
                class=C'
            RULE_LIST := RULE_LIST  $\cup$  r; {dodaj regułę r do RULE_LIST}

    until BEST_CPX = 0 or E =  $\emptyset$ .
return RULE_LIST.

Procedure Find_Best_Complex(E)
    Niech zbiór STAR zawiera tylko puste kompleksy;
    BEST_CPX := 0;
    Niech SELECTORS będzie zbiorem wszystkich możliwych selektorów;
    while STAR  $\neq$   $\emptyset$ ,
        specjalizacja kompleksów w zbiorze STAR:
        Niech NEWSTAR będzie zbiorem  $\{x \wedge y \mid x \in STAR, y \in SELECTORS\}$ ;
        Usuń wszystkie kompleksy w NEWSTAR które są również w
        STAR (tj. niewyspecjalizowane) lub są nieważne (np.
        big = y  $\wedge$  big = n)
        dla każdego kompleksu Ci w NEWSTAR:
            if Ci jest statystycznie znaczący testując na E i
            lepszy niż BEST_CPX zgodnie ze zdefiniowanymi
            przez użytkownika kryteriami testując na E,
            then BEST_CPX := Ci; {zamień bieżącą wartość
                BEST_CPX na Ci}
        repeat usuń najgorsze kompleksy z NEWSTAR
        until rozmiar NEWSTAR jest  $\leq$  zdefiniowane maximum;
        STAR := NEWSTAR;
return BEST_CPX.

```

Bramer [Bramer, 2000] proponuje usprawnioną wersję algorytmu *PRISM* zwaną *N-PRISM* pozwalającą uzyskać lepsze wyniki indukcji reguł z danych

zaszumianych. Innym algorytmem bazującym na sekwencyjnym pokrywaniu jest wspomniany już wcześniej algorytm *GTS*. Umożliwia indukcję reguł decyzji z niesprzecznych baz informacyjnych. Do oceny kandydatów na warunki elementarne stosuje się przedstawioną wyżej miarę H :

$$H = \text{Generality} + \text{sqrt}(\text{Accuracy}) \quad (3.2)$$

gdzie: $\text{Generality} = (E_p + E_b) / E$ oraz $\text{Accuracy} = E_p / E_p + E_b$; E_p – liczba poprawnie rozpoznawanych przypadków ze zbioru E , E_b – liczba błędnie rozpoznawanych przypadków ze zbioru E , E – liczba przypadków w zbiorze E ,

```

Procedure GTS ( $K_j$  - klasa;  $A$  - atrybuty;  $E$  - przykłady);
   $R := \emptyset$ ;                                {zbiór poszukiwanych reguł}
   $r := \text{learn-one-rule}(\text{klasa } K_j; \text{ atrybuty } A; \text{ przykłady } E)$ 
  while  $E \neq \emptyset$  do
     $R := R \cup r$ ;
     $E := E \setminus [R]$ ;          {usuń przykłady pozytywne pokryte przez  $R$ }
     $r := \text{learn-one-rule}(\text{klasa } K_j; \text{ atrybuty } A; \text{ przykłady } E)$ ;
  return  $R$ 

Procedure learn-one-rule(klasa  $K_j$ ; atrybuty  $A$ ; przykłady  $E$ );
 $E'$  - część warunkowa pierwszego przykładu ze zbioru  $E$ 
 $K'$  - klasa przykładu  $E'$ 
 $W' = \emptyset$ ;                                {kompleks warunków dla klasy  $K'$ }
while  $W' \neq E'$ 
  for  $w := W' \cup w_r$                         {dla każdej kombinacji  $W'$  z niewykorzystanymi warunkami  $w_r \in E' \setminus W'$ }
     $H(w) := \text{Generality} + \text{sqrt}(\text{Accuracy})$ ;
    if  $\text{Accuracy} = 1$ 
      then  $W' := w$ ; return  $r$ ; { $r$  jest regułą postaci IF  $W'$  THEN  $K'$ };
   $H_{\max} := \max\{H(w)\}$ ;                    {największa wartość  $H$  dla kompleksów  $w$ }
   $W' := w(H_{\max})$ ;                          {kompleks  $w(H_{\max})$  staje się trzonem tworzonej reguły}
return  $r$ ;                                  { $r$  jest regułą postaci IF  $W'$  THEN  $K'$ };

```

Jeden z bardziej znanych i skutecznych algorytmów indukcji reguł zaproponował Grzymała-Busse [Grzymała-Busse, 1992]. Algorytm *LEM2* pozwala generować reguły również w przypadku zbiorów sprzecznych. Wykorzystuje w tym celu teorię zbiorów przybliżonych [Pawlak, 1982; Pawlak, 1991]. Algorytm *LEM2* tworzy minimalny opis dyskryminujący przybliżenia danej klasy decyzyjnej za pomocą specyficznej zasady generowania kolejnych pokryć. Algorytm ten został zaimplementowany w systemie uczenia maszynowego *LEERS* (*Learning from Examples based on Rough*

Sets) [Grzymała-Busse, 1992; Grzymała-Busse, 1997]. Ogólną postać algorytmu *LEM2* przedstawiono poniżej.

```

Procedure LEM2
input:  $K$ ; {zbiór przykładów pozytywnej klasy  $K_j$ }
output:  $R$ ; {zbiór reguł opisujących klasę  $K_j$ }
begin
   $G := K$ ; { $G$  (ang. goal czyli cel) zbiór obiektów nie pokrytych
             dotychczas przez elementy z  $\mathbf{P}$ }
   $\mathbf{P} := \emptyset$ ; {lokalne pokrycie  $\mathbf{P}$  zbioru  $K$ }
  while  $G \neq \emptyset$ 
    begin
       $P := \emptyset$ ; {kandydat na część warunkową reguły}
       $W(G) := \{w: [w] \cap G \neq \emptyset\}$ ; {zbiór potencjalnych wa-
                                             runków elementarnych}

      while  $P = \emptyset$  or not  $[P] \subseteq K$  {iteracja ta jest przykła-
                                             dem funkcji learn-one-rule}
        begin
          wybierz warunek  $w \in W(G)$  posiadający atry-
          but o największym priorytecie, jeżeli nie
          ma takiego atrybutu wybierz warunek  $w \in$ 
           $W(G)$  taki, że wyrażenie  $|[w] \cap G|$  ma naj-
          większą wartość, jeżeli więcej niż jeden z
          warunków posiada maksymalną wartość powyż-
          szego wyrażenia, to wybierz ten, dla któ-
          rego wyrażenie  $|[w]|$  ma najmniejszą war-
          tość, w przypadku równości wybierz pierw-
          szy z rozważanych warunków

           $P := P \cup \{w\}$ ; {dołącz najlepszy warunek
                              do koniunkcji  $P$ }

           $G := [w] \cap G$ ; {ogranicz zbiór obiektów
                              dla stworzenia warunków elementarnych}

           $W(G) := \{w : [w] \cap G \neq \emptyset\}$ ; {uaktualnij
                              listę potencjalnych warunków}

           $W(G) := W(G) - \{P\}$ ;

        end {while}

      for  $w \in P$  do {dla każdego warunku elementarnego w
                          należącego do koniunkcji  $P$ }
        if  $[P - \{w\}] \subseteq K$  then  $P := P - \{w\}$ ; {usuń warunki
                          nadmiarowe}

       $P := P \cup \{P\}$ ;
       $G := K - \cup_{P \in \mathbf{P}} [P]$ ;
    end {while};

    for  $P \in \mathbf{P}$  do {dla każdego kompleksu  $P$  należącego do po-
                          krycia  $\mathbf{P}$ }
      if  $\cup_{T \in \mathbf{P} - \{P\}} [T] = K$  then  $\mathbf{P} := \mathbf{P} - \{P\}$ ; {usuń ewentu-
                          alne nadmiarowe reguły}
    return  $R$ ; {utwórz zbiór reguł  $R$  na podstawie koniunkcji z  $\mathbf{P}$ }
  end.

```


Procedura *LEM2* uruchamiana jest iteracyjnie dla każdej z klas decyzyjnych. Jest tu zastosowana pewna notacja: $[w]$ oznacza zbiór przykładów spełniających warunek w , $W(G)$ oznacza zbiór warunków elementarnych, które można zbudować na podstawie opisów przykładów ze zbioru $G \subseteq K$ (według kolejności wartości atrybutów występujących w opisie przykładów z G), P jest koniunkcją (kompleksem) warunków elementarnych, która jest kandydatem na część warunkową reguł, ponadto P powinno być minimalną koniunkcją tzn. usunięcie któregoś z warunków w_i prowadzi do $[P] \not\subseteq K$. Kompleks P konstruowany jest z użyciem warunków $w \in W(G)$ do momentu aż $[P] \subseteq K$ co jest końcowym kryterium budowy reguły w algorytmie *LEM2*. W stosowanej notacji, \mathbf{P} jest zbiorem koniunkcji P stanowiącym tzw. *lokalne pokrycie* (ang. *local covering*) zbioru przypadków klasy K . Na podstawie tychże koniunkcji tworzone są reguły decyzji wskazujące właściwe klasy decyzyjne.

W późniejszych badaniach można odnaleźć wiele modyfikacji algorytmu *LEM2*: algorytm *MLEM2* [Grzymała-Busse, 2002; Grzymała-Busse, 2003, Grzymała-Busse, 2003a], który wykonuje dyskretyzację atrybutów numerycznych równocześnie z procesem indukcji reguł. Podobnie jest w przypadku algorytmu *MODLEM* [Stefanowski, 1998], gdzie tworzenie uogólnionych warunków elementarnych obejmujących przedziały wartości jest wykonywane lokalnie w trakcie indukcji reguł. Z kolei w algorytmie *ELEM2* [An i Wang, 2001; An, 2003] indukcja reguł decyzji opiera się na heurystycznym przeszukiwaniu przestrzeni hipotez. Algorytm ten zawiera nową funkcję heurystyczną kierującą tym procesem oraz analizuje dane sprzeczne poprzez zdefiniowanie obszarów niemożliwych do poprawnego nauczania. Obszary takie stanowią kryterium kończące proces indukcji reguł w danej klasie.

Oprócz algorytmów generujących reguły *bezpośrednio* przy użyciu tablicy decyzyjnej, istnieją również *pośrednie* metody generowania reguł

decyzji przekształcające inne formalizmy reprezentacji wiedzy na zbiór reguł decyzji. Z tej grupy najbardziej rozpowszechnione są metody konwersji drzew decyzji w zbiór reguł decyzji. Ten sposób został zaproponowany przez R. Quinlana w algorytmie budowy drzew decyzji *C4.5* (rozbudowana wersja algorytmu *ID3*) [Quinlan, 1993; Quinlan, 1996], który polega na odwzorowaniu poszczególnych ścieżek w drzewie, od korzenia do liści do postaci reguł JEŻELI ... TO... . W metodzie generowania reguł decyzji z sieci przekonań [Hippe, Grzymała-Busse i Mroczek, 2004], sieć jest zamieniana na zbiór reguł przy uwzględnieniu wielkości wpływu poszczególnych atrybutów na decyzję. Szerzej znane są również metody ekstrakcji reguł przy użyciu algorytmów generujących sieci neuronowe [Duch, Adamczak i Grąbczewski, 1996; Duch, Adamczak i Grąbczewski, 1999] czy algorytmy ewolucyjne rozwijające dotychczasowe możliwości znanych algorytmów indukcji reguł [Pappa i Freitas, 2004; Kwedło i Krętowski, 2002]. Istnieją także metody generowania reguł decyzji poprzez zastosowanie algorytmów analizy skupień (ang. *cluster analysis*). Jednym z algorytmów tego typu jest *PNC2* (*Positive and Negative example-based Clustering*) [Haendel, 2003], w którym poszczególne pary danych (wielowymiarowy wektor wejściowy oraz odpowiadająca mu wartość wyjściowa) są traktowane jako oddzielne skupiska. Skupiska o tych samych wartościach wyjściowych są grupowane. W następnym etapie, w każdej z grup dwa najbliższe sobie skupiska, wcześniej nie rozpatrywane, łączone są iteracyjnie do momentu osiągnięcia tzw. *kryterium aborcji*, lub do momentu gdy wszystkie pary danych połączone są w jedno skupisko. Przed połączeniem dwóch skupisk *PNC2* dokonuje testu czy połączone skupisko będzie stanowiło prawidłową regułę.

Znane są również podejścia hybrydowe do generowania reguł, np. hybrydowe połączenie drzew decyzji oraz algorytmów genetycznych *C4.5/GA* [Carvalho i Freitas, 2004]. Takie podejście składa się z dwóch etapów. W pierwszym etapie tworzone jest przy pomocy algorytmu *C4.5*

przycięte drzewo decyzji i zamieniane jest na zbiór reguł decyzji. Każda z utworzonych reguł jest rozpatrywana jako *mała* lub *duża separowalność* (ang. *disjunct*) w zależności od tego, czy liczba przypadków pokrywanych przez regułę jest mniejsza czy równa zadanemu parametrowi progowemu. W drugim etapie algorytm genetyczny znajduje reguły pokrywające przypadki małej separowalności. Przypadki przynależne do dużej separowalności są klasyfikowane przez drzewo decyzji utworzone przez C4.5.

Stosowanie reguł decyzji do klasyfikowania nowych obiektów

Reguły decyzji wygenerowane ze zbioru przykładów uczących używane są do klasyfikowania nowych, nieznanych obiektów. Przez nowe obiekty rozumie się obiekty, które nie były wykorzystywane w procesie indukcji reguł. Znany jest ich opis za pomocą wartości atrybutów, natomiast celem klasyfikowania jest określenie przydziału takiego obiektu do klasy decyzyjnej. Jeżeli dla klasyfikowanego obiektu znana jest ponadto jego rzeczywista klasyfikacja to nazywamy go przykładem testowym, gdyż możliwe jest wtedy porównanie proponowanej decyzji klasyfikacyjnej z rzeczywistą. Klasyfikowanie obiektów opiera się na dopasowaniu opisu obiektu do części warunkowych reguł decyzji. Wiele strategii klasyfikacji dla reguł decyzji zostało zaproponowanych w ostatnich dwóch dekadach, wiele z nich jest rozważanych w [Stefanowski, 2001; Wróblewski, 2001]. Można je z grubsza podzielić na: *i*) metody klasyfikujące w oparciu o listę reguł (reguły są wykorzystywane zgodnie z kolejnością występowania na liście) oraz *ii*) metody wykorzystujące zbiory reguł (kolejność reguł nie ma znaczenia) [Pindur, 2004]. W podstawowej formie proces klasyfikacji obiektów przez reguły można podzielić na etap poszukiwania reguł, które odpowiadają obiektowi e oraz etap przyporządkowania obiektu klasie wyznaczonej przez obiekt. W kontekście dopasowania reguły do badanego obiektu wyróżnia-

my *dopasowanie całkowite* i *częściowe*. Niech R oznacza zbiór reguł uzyskany w wyniku indukcji oraz niech $R_K \subset R$ oznacza podzbiór reguł odpowiadający pojedynczej klasie decyzyjnej K_K . Reguły są reprezentowane w najogólniejszej postaci (2.3). Rozpatrzmy zbiór nowych obiektów E^n . Niech $e \in E^n$ oznacza obiekt, który ma zostać sklasyfikowany. Jego opis można zapisać jako $\bigwedge_{a_i \in A} (f(a_i, e) = v_{ie})$, gdzie v_{ie} jest wartością atrybutu a_i z dziedziny V_{a_i} . Zbiór atrybutów wykorzystany do skonstruowania części warunkowej reguły r oznaczmy przez C_r .

Mówimy o **całkowitym dopasowaniu** (ang. *complete matching*) obiektu e do części warunkowej reguły r , jeśli wszystkie atrybuty a_i obiektu e zawierają się w zbiorze C_r . Mówimy o **częściowym dopasowaniu** (ang. *partial matching*) obiektu e do części warunkowej reguły r , jeśli istnieje przynajmniej jeden atrybut $a_j \in C_r$, dla którego nie istnieje atrybut tożsamy (równy) w zbiorze atrybutów a_i obiektu e . Dopasowanie reguł decyzji do opisu klasyfikowanego przykładu realizowane jest różnie w zależności od tego, czy reguły decyzji uporządkowane są w postaci listy decyzyjnej, czy tworzą nieuporządkowany zbiór reguł. W przypadku listy decyzyjnej dokonuje się dopasowania obiektu do kolejnych reguł na liście [Cestnik, 1990; Quinlan, 1993]. Pierwsza dopasowana reguła na liście wyznacza przydział obiektu e do klasy decyzyjnej. Ostatnią zaś regułą na liście jest tzw. *reguła domyślna* (ang. *default rule*), stosowana wtedy, gdy żadna z poprzednich reguł nie dopasowała się do obiektu. Reguła domyślna najczęściej wskazuje klasę większościową w zbiorze uczącym, tj. najliczniejszą klasę w zbiorze uczącym. W przypadku klasyfikowania obiektu za pomocą nieuporządkowanego zbioru reguł występuje jedna z następujących sytuacji: **(i)** część warunkowa dokładnie jednej reguły jest całkowicie dopasowana do obiektu, który jest zaklasyfikowany do klasy wskazywanej przez tę regułę, **(ii)** część warunkowa więcej niż jednej reguły jest całkowicie dopasowana do obiektu, oraz **(iii)** część warunkowa żadnej z reguł nie jest dopasowana do obiektu.

W pierwszym przypadku sytuacja jest jednoznaczna i nie wymaga dodatkowych objaśnień, natomiast w pozostałych przypadkach stosuje się różne rozwiązania [Bazan, 1998; Clark i Boswell, 1991; Grzymała-Busse, 1994; Słowiński i Stefanowski, 1994; Michalski, Carbonell i Mitchell, 1986]. Jednym z najbardziej znanych jest podejście Grzymały-Busse, który proponuje rozwiązanie obu sytuacji na podstawie dodatkowych miar charakteryzujących regułę r . Gdy ma miejsce sytuacja druga obliczane jest *poparcie* dla danej klasy decyzyjnej:

$$SUP(K_K) = \sum_{r \in R_K^-} Sila(r) * Liczba_warunków(r) \quad (3.3)$$

gdzie $Sila(r)$ jest liczbą poprawnie klasyfikowanych przypadków ze zbioru uczącego przez regułę r . R_K^- oznacza reguły z R_K (wskazujące klasę K_K) dopasowane do obiektu nieznanego. Przypadek nieznaną jest przydzielany do klasy K_K , dla której poparcie $SUP(K_K)$ ma największą wartość. W sytuacji trzeciej należy najpierw zidentyfikować reguły częściowo dopasowane do obiektu, a następnie określić dodatkową miarę *stopnia dopasowania* reguły r do obiektu e , tj. $MF(r, e)$ zwany również *matching factor*. Miara ta jest równa stosunkowi liczby warunków reguły dokładnie dopasowanych do opisu obiektu do ogólnej liczby warunków tworzących część warunkową reguły.

$$MF(r, e) = \frac{liczba_warunków_dopasowanych_reguly_r}{liczba_wszystkich_warunków_w_regule_r} \quad (3.4)$$

Dla reguł częściowo dopasowanych z danej klasy K_K *poparcie* jest równe wówczas

$$SUPP(K_K) = \sum_{r \in R_K^-} MF(r, e) * Sila(r) * Liczba_warunków(r) \quad (3.5)$$

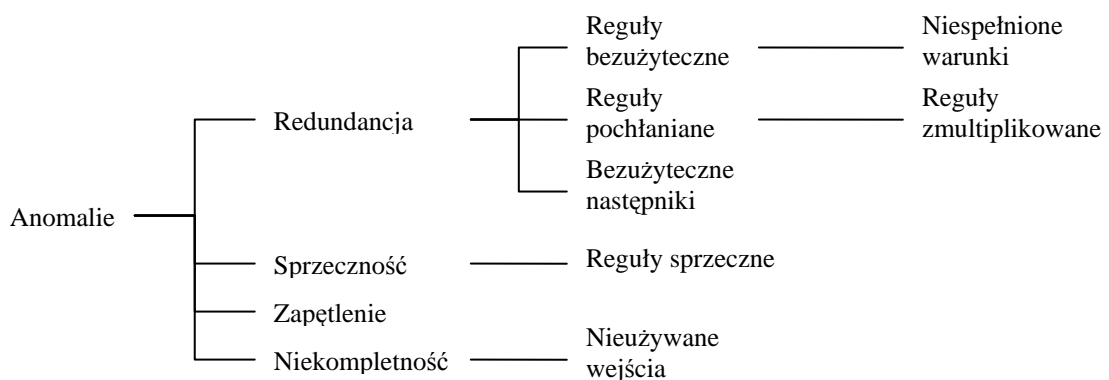
gdzie R_K^- oznacza reguły R_K częściowo dopasowane do obiektu e . Obiekt e jest przypisywany do klasy K_K która ma największą wartość $SUPP(K_K)$. Przedstawiona metoda klasyfikacji obiektów nieznaną jest również sto-

sowana w dalszej części rozprawy. Algorytmy indukcji reguł decyzji mogą powodować powstawanie pewnych anomalii w wygenerowanych zbiorach reguł. Zagadnienie to jest tematem kolejnego rozdziału poniższej rozprawy.

ROZDZIAŁ 4

WERYFIKACJA REGUŁOWEJ BAZY WIEDZY

W poprzednim rozdziale omówiono ogólne cechy procesu indukcji reguł decyzji oraz wybrane algorytmy generowania reguł. Różnorodność algorytmów, a także charakter danych w tablicy decyzyjnej mogą powodować powstawanie pewnych anomalii, spotykanych w modelach uczenia w postaci zbiorów reguł. Proces *weryfikacji* zbiorów reguł pozwala ustalić ich logiczną poprawność, poprzez sprawdzenie kontekstu [Preece i Shinghal, 1994; Spreeuwenberg i Gerrits, 2002] logicznej konsystencji, występowania zapętleń oraz kompletności, redundancji i aktualności reguł. Preece sformułował pojęcia anomalii spotykanych w bazach reguł, według schematu przedstawionego na Rys. 4-1.



Rys. 4-1 Cztery typy anomalii sformułowanych wg [Preece i Shinghal, 1994]

Powyższe pojęcia w późniejszym okresie podlegały ewolucji [Harmelen, 1998]. *Weryfikacja* nie powinna być mylona z technikami *walidacji*, co

przedstawił Gonzales [Gonzales i Barr, 2000], omawiając kontrowersje w definiowaniu tego pojęcia. Walidacja pozwala ustalić poprawność systemu w aspekcie użycia go w określonej dziedzinie i środowisku. Zgodnie z [Spreeuwenberg i Gerrits, 2002] uznaje się, że walidacja jest określana jako „budowanie poprawnego produktu”, natomiast weryfikacja to „budowanie produktu poprawnie”. Można również stwierdzić, iż weryfikacja jest wymagana przez walidację i jest jej pod-zadaniem [Laurent, 1992]. Narzędzia związane z weryfikacją i walidacją regułowej bazy wiedzy są nazywane w literaturze narzędziami V&V.

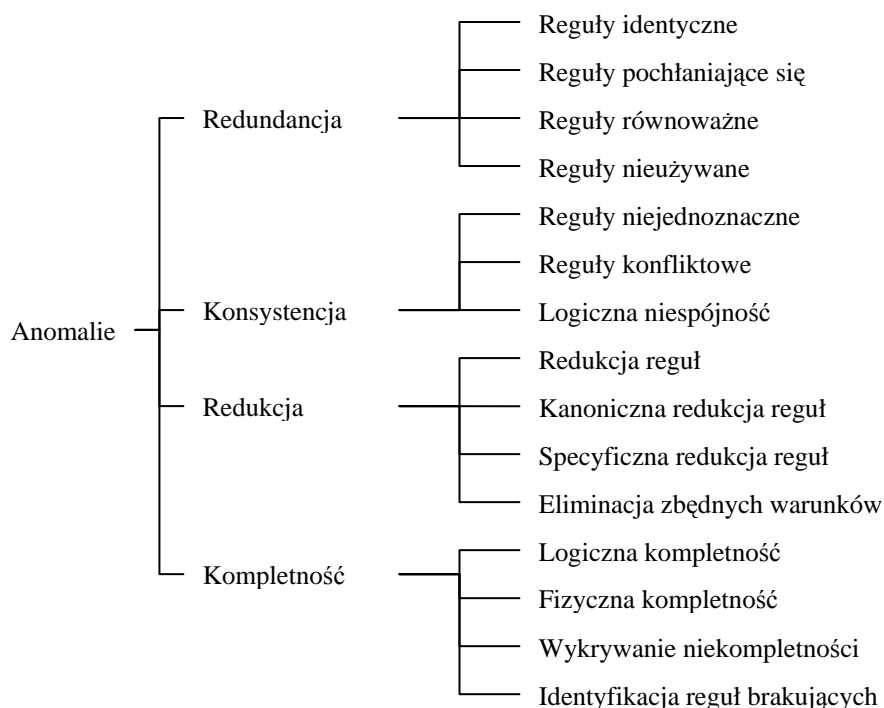
Na początku lat 90-tych, różne grupy badawcze pracowały nad rozwojem systemów V&V. Powstały narzędzia weryfikujące bazy reguł [Preece i Shinghal, 1994]. Narzędzia te zostały dokładniej omówione w [Murrell i Plant, 1995], którzy scharakteryzowali 35 narzędzi typu V&V zbudowanych w okresie 1985-1995. Na początku lat 80-tych, systemy tego typu wykrywały anomalie związane z wiedzą nadmiarową (zmultiplikowane i pochłaniające się pary reguł), reguły sprzeczne (pary reguł z równoważnymi warunkami lecz ze sprzecznymi konkluzjami) i reguły brakujące (logiczne kombinacje dozwolonych danych wejściowych, nie pokrywające warunków w istniejących już regułach). Należy do nich zaliczyć systemy *Rule Checker Program (RCP)* [Suwa, Scott i Shortliffe, 1982], *CHECK* [Nguyen i inni, 1985], czy systemy zajmujące się zagadnieniem sprawności, np. *ESC* [Cragun i Steudel, 1987]. Nieco później na początku lat 90-tych, znaczenie systemów detekcji anomalii wzrosło poprzez szersze zdefiniowanie anomalii poza prostymi parami reguł: redundancje i sprzeczności pojawiające się w łańcuchach reguł mogą być wykryte również w subtelnych przypadkach brakującej wiedzy. Przykładem mogą być systemy *COVADIS* [Rousset, 1988], *KB-Reducer* [Ginsberg, 1988], *COVER* [Preece, 1989], który ewoluował później do postaci *COVERAGE* (weryfikacja baz reguł w architekturze multiagentowej) [Preece, 1999], *VALENS (VALid ENgineering Support)* [Gerrits, Spreeuwenberg i Boekenoogen, 2000], stosujący zarówno techniki

weryfikacji jak i walidacji bazy reguł. Niektóre algorytmy generowania reguł decyzji posiadają wbudowane niektóre operacje weryfikujące generowany zbiór reguł. Przykładem może być algorytm *LEM2* zaimplementowany w systemie *LERS*, przedstawiony w poprzednim rozdziale. Algorytm ten zawiera procedury odpowiedzialne za usunięcie nadmiarowych warunków i reguł ze zbioru reguł. Innym przykładem systemu weryfikującego wiedzę w postaci reguł decyzji jest system *CILT* (ang. *Cooperative Induction Learning Team*). Jest to hybrydowy multi-strategiczny system łączący kilka technik generowania reguł decyzji na podstawie tego samego zbioru danych. *CILT* umożliwia utworzenie reguł przy pomocy algorytmu generowania drzew decyzji *C4.5* [Quinlan, 1993], algorytmu *CN2* [Clark i Niblett, 1989] generującego reguły metodą kolejnych pokryć oraz algorytm *ANN-SER* agregujący reguły decyzji z nauczonej sieci neuronowej [Viktor, Engelbrecht i Cloete, 1998]. W systemie tym indywidualne rezultaty poszczególnych metod są łączone w jeden w trójfazowym procesie [Viktor, 1999; Viktor, Paquet i le Roux, 2003]: (i) *uczenie indywidualne* – wstępne modele uczenia w formie reguł są tworzone przez poszczególne algorytmy i przechowywane są jako oddzielne komponenty bazy wiedzy, (ii) *uczenie wspólne* – poszczególne modele wstępne dzielą się wiedzą nawzajem. Do odróżnienia reguł o większej i mniejszej jakości używany jest próg dokładności (ang. *accuracy threshold*). Uczenie to składa się z czterech kroków w których budowany jest zestaw Nowych reguł decyzji zawierający najbardziej wartościowe spośród wygenerowanych przez poszczególne algorytmy, następnie relacje pomiędzy regułami w uzyskanym zbiorze podlegają identyfikacji i kompilacji. System rozpoznaje podobieństwa reguł (ang. *similarities*), reguły zachodzące (ang. *overlapping rules*) oraz pochłanianie się reguł (ang. *subsumption*). W kroku trzecim wykonywana jest procedura kombinacji reguł. Reguły zawarte w Nowym zbiorze są używane do uformowania nowych reguł następująco: pary atrybut-wartość reguł z Nowego zbioru są kombinowane z takimi parami zawartymi w zbiorach wygenero-

wanych przez użyte narzędzia. Każda z formowanych reguł jest testowana a reguły wysokiej jakości, są różne od siebie, nie nachodzą na siebie, i nie są pochłaniane przez istniejące reguły pozostają w Nowym zbiorze. Służą one jako dane wejściowe do kroku czwartego jakim jest generator danych. Każda reguła pozostająca w Nowym zbiorze służy do utworzenia nowych przypadków uczących, które następnie dodawane są do oryginalnego zbioru danych. Cały proces rozpoczyna się od nowa. Wszystkie kroki powtarzają się iteracyjnie aż do momentu gdy nie można utworzyć już żadnej nowej reguły. (iii) *fuzja wiedzy* – wiedza zawarta w każdej z baz regułowych jest łączona w jedną. Koncepcja działania systemu *CILT* ma podobny charakter jak proponowane w rozprawie podejście budowy optymalnych modeli uczenia. Zagadnienie to zostanie jednak poruszone w dalszych rozdziałach szczegółowo przedstawiających proponowaną metodologię.

W odniesieniu do regułowej reprezentacji wiedzy, zagadnienia związane z anomaliami można pogrupować następująco [Ligęza, 2005]: (i) *redundancja*: reguły identyczne (ang. *identical rules*), reguły pochłaniające się (ang. *subsumed rules*), reguły równoważne (ang. *equivalent rules*), reguły nieużywane (ang. *unusable rules*), (ii) *konsystencja (spójność)*: reguły niejednoznaczne (ang. *ambiguous rules*), reguły z konfliktem (ang. *conflict rules*), reguły z niespójnością logiczną, (iii) *redukcja*: redukcja reguł, kanoniczna redukcja reguł, specyficzna redukcja reguł, eliminacja zbędnych atrybutów, (iv) *kompletność*: logiczna kompletność, fizyczna kompletność, wykrywanie niekompletności, identyfikacja reguł brakujących.

Należy zaznaczyć, że powyższa klasyfikacja nie odnosi się do problemów wymagających głębszej analizy np. sprzeczności reguł, ich zapętlenia, nieosiągalnych konkluzji, itp. Proponowana na Rys. 4-2 klasyfikacja jest dość ogólna, jednak zawiera wiele szczególnych przypadków opisywanych w literaturze. Poniżej omówiono pokrótce poszczególne grupy anomalii w regułach.



Rys. 4-2 Cztery typy anomalii regułowej reprezentacji wiedzy wg [Ligęza, 2005]

Redundancja reprezentacji wiedzy

Redundancja reprezentacji wiedzy dotyczy problemu oceny skuteczności opisu elementów wiedzy, uwolnionego od niepotrzebnych składników. Zagadnienie to w większości przypadków nie ma wpływu na poprawność i funkcjonalność działania systemów regułowych, jednak powoduje spowolnienie działania procesu wnioskowania i może się stać źródłem trudności podczas modyfikacji i aktualizacji bazy reguł. Ponadto, elementy redundantne zajmują pamięć i komplikują analizę reguł. Podstawową metodą pozbycia się redundancji wiedzy jest wykrycie i usunięcie składników powtarzających się. Ligęza [Ligęza, 2005] wyróżnia dwa rodzaje redundancji: *logiczną*, tzn. taką która dotyczy powtarzalności reguł w sensie ich budowy logicznej i *funkcjonalną*, dotyczącą reguł, które mogą zostać usunięte ze zbioru bez wpływu na jego funkcjonowanie. Redundancja funkcjonalna niejako wynika z redundancji logicznej. Rzeczą oczywistą jest to, że powta-

rzające się, identyczne reguły powinny być wyeliminowane ze zbioru reguł, podobnie jak reguły równoważne. Najbardziej interesującym przypadkiem są jednak reguły powtarzające się. Tego typu reguły mogą być również usuwane ze zbioru reguł decyzji bez wpływu na jego rolę w procesie wnioskowania. W pewnych przypadkach jednak pozostawienie pochłanianej, bardziej szczegółowej reguły, może być celowe, aby nie powodować konfliktów i nie wpływać na strategię wnioskowania [Lunardi i Passino, 1995]. Równoważność reguł dotyczy reguł które nawzajem się pochłaniają; jest to jeden ze specyficznych przypadków reguł pochłaniających się. Dwie reguły są równoważne, jeżeli jedna pochłania drugą i odwrotnie. W związku z tym, pochłanianie jest zagadnieniem zawierającym w sobie problematykę reguł identycznych oraz równoważnych. Natomiast reguły nieużywane dotyczą innego zagadnienia: reguły takie istnieją, jeżeli co najmniej jeden warunek w regule nie jest nigdy spełniony.

Niedeterministyczność i niekonsystencja reguł decyzji

Zagadnienie niedeterministyczności (niejednoznaczności) i niekonsystencji reguł odnosi się do problemu wewnętrznej niespójności regułowej bazy wiedzy. W takim przypadku, zastosowanie reguł może prowadzić do niejednoznacznych i niespójnych rezultatów. W tym kontekście można wyróżnić następujące anomalie: (i) *niedeterministyczne, niejednoznaczne reguły*, (ii) reguły *ambivalentne, konfliktowe*, oraz (iii) *logiczną niekonsystencję*.

Mała unikalność reguł może prowadzić do *niejednoznacznych konkluzji*. Innymi słowy, niejednoznaczność powstaje w przypadku gdy dwie (lub więcej) reguł mogą być zastosowane dla tych samych wartości wejściowych, ale ich konkluzja jest różna. Bardziej kłopotliwy przypadek stanowią reguły, powodujące konflikty. Ma to miejsce kiedy dwie, współbieżnie uzyskane decyzje, nie mogą być poprawne w związku z zamierzoną inter-

pretacją (światem zewnętrznym). Innymi słowy, reguły konfliktowe mogą być zastosowane równocześnie, ale ich konkluzje powodują konflikt, nie mogą być równocześnie prawdziwe. Konflikt może również wywołać *logiczna niekonsystencja*, tzn. przypadek gdy reguły (konfliktowe) mogą być równocześnie zastosowane, ale ich konkluzje są logicznie niekonsystentne, na przykład jedna jest zaprzeczeniem drugiej, co nie może być prawdziwe w kontekście żadnej interpretacji. Charakterystyka i źródło powyższych problemów dotyczy problematyki zachodzących za siebie warunków w regułach. W związku z tym podstawowa procedura weryfikacji powinna umożliwić odkrycie reguł, których warunki mogą być równocześnie spełnione.

Podsumowując, zbiór reguł jest *deterministyczny*, gdy nie istnieją dwie różne reguły możliwe do zastosowania dla tego samego stanu wejściowego. *Niedeterministyczny* zbiór reguł nazywany jest również zbiorem *niejednoznaczny*. W przypadku zbioru niedeterministycznego pojawia się problem odpowiedniego wyboru jednej reguły, która będzie zastosowana do klasyfikacji przypadku nieznanego. Najczęściej stosowana jest odpowiednia miara pozwalająca wybrać najlepszą regułę. Wybrane miary oceny reguł zostały przedstawione w poprzednim rozdziale.

Redukcja zbioru reguł decyzji

Problemy minimalnej, w domyśle optymalnej reprezentacji wiedzy, odnoszą się do możliwości transformacji początkowej wiedzy do możliwie najprostszej postaci. Można to uzyskać poprzez *redukcję wiedzy*, jednakże jej zredukowana postać, powinna być logicznie równoważna wiedzy pierwotnej. Dlatego redukcja wiedzy jest rozumiana nie jako proste usunięcie zbędnych reguł (np. pochłaniających się). W poszukiwaniu minimalnej reprezentacji reguł, dwie lub więcej reguł mogą być łączone (*sklejane*, ang. *gluing operation*) w jedną,

wych. W redukcji reguł jest niezbędnym, aby wartości atrybutów pozostających w przeciwieństwie do atrybutów redukowanych były identyczne.

Redukcja częściowa reguł może prowadzić do specyficznej, unikalnej postaci, zwanej *kanoniczną*. Rezultat redukcji całkowitej może nie być unikalny. W pewnych przypadkach skuteczniejsze może być budowanie nowej, bardziej ogólnej reguły, z uwzględnieniem wybranych atrybutów (tj. łącząc częściowo ich wartości), czy dopuszczenie nieznaczących ograniczeń dotyczących pozostałych warunków. Można zastosować następujący schemat redukcji:

$$\begin{array}{rcl}
 r^1: \varphi^1 \wedge \omega_1 & \rightarrow & h \\
 r^2: \varphi^2 \wedge \omega_2 & \rightarrow & h \\
 \vdots & & \vdots \\
 r^k: \varphi^k \wedge \omega_k & \rightarrow & h \\
 \hline
 r : \varphi^1 \wedge \varphi^2 \wedge \dots \wedge \varphi^k & \rightarrow & h
 \end{array}$$

Schemat 4-2

Może to być spełnione gdy formuła $\varphi^1 \wedge \varphi^2 \wedge \dots \wedge \varphi^k$ jest użyteczna, i może być stosowana przy weryfikacji kompletności. Należy zauważyć, że pełna logiczna równoważność może nie być zachowana, jednakże dysjunkcja warunków reguł początkowych wynika z warunków reguł zredukowanych. Jeżeli wynikowa reguła zredukowana może być zastosowana, to co najmniej jedna reguła składowa (Schemat 4-2) może być również zastosowana, a osiągnięte konkluzje są oczywiście identyczne.

Szczegółowa redukcja reguł polega na zamianie dwóch lub więcej wartości, wyrażeń czy też wzorców na inny, bardziej ogólny. Standardowo, redukcja prowadzi do eliminacji zbędnych wyrażeń i wzorców, natomiast w przypadku szczegółowej częściowej redukcji, elementy te pozostają ale w formie bardziej ogólnej. W pewnych przypadkach redukcja całkowita nie może zostać zrealizowana, jednak nadal możliwe jest uproszczenie zbioru reguł, gdy podformuła ω_i , $i=1,2,\dots,k$ może być zastąpiona przez pojedynczą równoważną formułę. Zestaw pewnych elementów może być zawsze zastąpiony przez podzbiór za-

ta) *kompletność* reguł oznacza, że dysjunkcja warunków wszystkich reguł tworzy tautologię, tj. nieważne jakie dane podane są na wejście, będą obsługiwane. Zbiór reguł postaci:

$$\begin{array}{l} r_1: \varphi_1 \rightarrow h_1 \\ r_2: \varphi_2 \rightarrow h_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ r_m: \varphi_m \rightarrow h_m \end{array}$$

Schemat 4-4

jest logicznie kompletny gdy $\models \varphi_1 \vee \varphi_2 \vee \dots \vee \varphi_m$, tzn. jeżeli zbiór warunków jest tautologią. Logiczna kompletność mówi, że bez względu na to co podane jest na wejście systemu regułowego, warunki co najmniej jednej reguły są spełnione. Dalej, co najmniej jedna reguła jest odpalona. W praktyce nie cały zestaw wejść może być dopuszczalny, lub zbiór reguł może być utworzony do pracy tylko z pewnymi wartościami wejściowymi (w specyficznym kontekście). Dla większości praktycznych systemów regułowych logiczna kompletność narzuca zbyt restrykcyjne wymagania. *Specyficzna (częściowa, fizyczna) kompletność* oznacza, że zakres wejść jakie system jest w stanie obsłużyć jest wyraźnie zdefiniowany przez formułę (ograniczającą warunki) określającą dopuszczalną przestrzeń wejść. Zatem zbiór reguł $R = \{r_1, r_2, \dots, r_m\}$ jest *specyficznie (fizycznie) kompletny* odpowiednio do rozpatrywanego kontekstu Ψ , jeżeli $\Psi \models \varphi_1 \vee \varphi_2 \vee \dots \vee \varphi_m$, tzn. jeżeli zbiór warunków jest tautologią w ramach kontekstu Ψ . Ten rodzaj kompletności jest bliższy praktycznemu rozumieniu kompletności. Jeżeli wszystkie fizyczne stany wejściowe które powinny być obsługiwane przez reguły są zdefiniowane formułą Ψ , wtedy specyficzna kompletność w kontekście Ψ oznacza również fizyczną kompletność pod względem wszystkich stanów spełniających Ψ .

Jeżeli zbiór reguł nie spełnia wymagań kompletności, interesującym staje się określenie luk w zbiorze tj. *detekcja niekompletności* i późniejsza *identyfikacja reguł brakujących*.

Kolejne rozdziały będą stanowić część badawczo-doświadczalną rozprawy. Na podstawie powyższych informacji teoretycznych zostanie przedstawiona plan i metodyka proponowanych badań, analiza wyników oraz wnioski.

ROZDZIAŁ 5

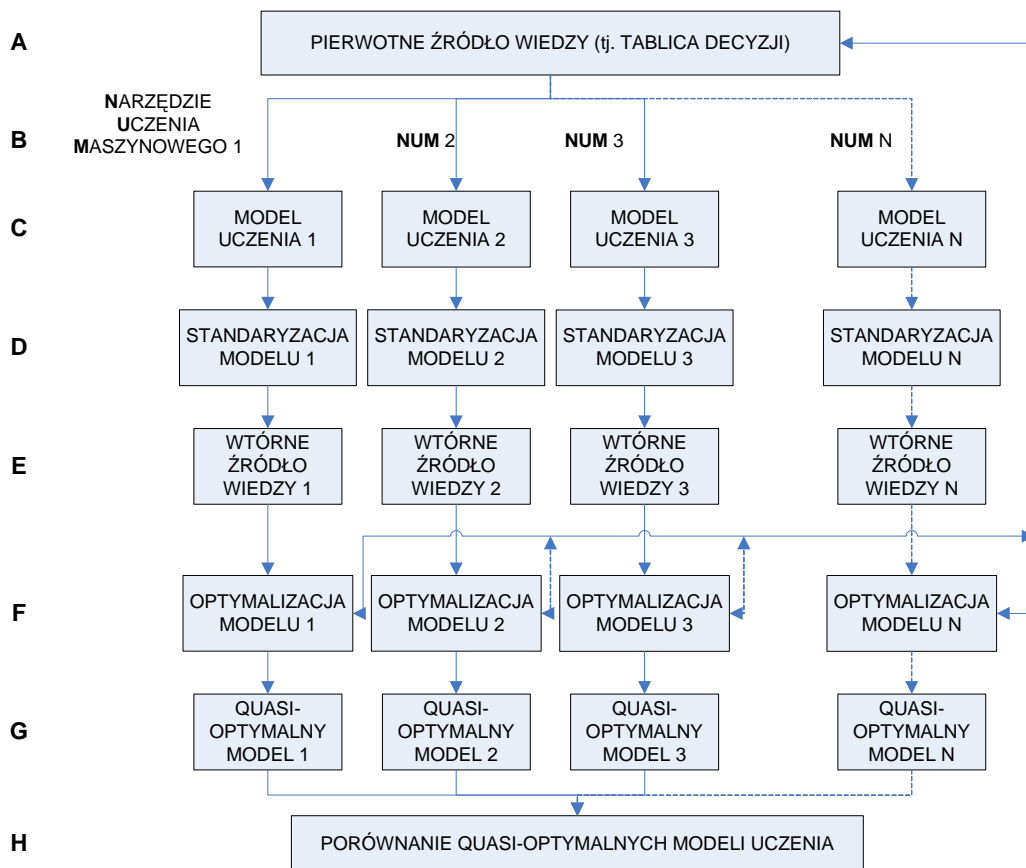
KONCEPCJA OPTIMALIZACJI WTÓRNYCH ŹRÓDEŁ WIEDZY

Modele uczenia pozyskiwane w procesie indukcji, powstają na podstawie danych historycznych, i są stosowane do objaśnienia danych nieznanymi. W relacji do jakości modeli uczenia, znany paradygmat sztucznej inteligencji stwierdza, że *zastosowanie różnych narzędzi nadzorowanego uczenia maszynowego do analizy danego (tego samego) zbioru danych, prowadzi zazwyczaj do utworzenia odmiennych modeli uczenia, objaśniających z różnymi skutkami i z różną precyzją ukrytą w nim wiedzę* [Winston, 1992]. Modele te mogą stosować różne formalizmy reprezentacji wiedzy; spośród szerokiego spektrum tych formalizmów, najbliższym zapisowi wiedzy stosowanemu przez człowieka, jak wspomniano w Rozdziale 2, są reguły decyzji [Cichosz, 2000; Michalski i inni, 1986; Mitchell, 1997; Stefanowski, 2001]. Są one indukowane na podstawie zbiorów danych, reprezentują regularności i zależności występujące w opisie *obiektów*, nazywanych również *przykładami uczącymi*. W większości przypadków algorytmy indukcji reguł decyzji (szerzej przedstawione w Rozdziale 3) ukierunkowane są na klasyfikację nowych obiektów oraz generowanie minimalnych zbiorów reguł, opisujących jak największą liczbę przykładów uczących. Zazwyczaj główną wadą wspomnianych algorytmów jest zdolność generowania zbyt dużej liczby reguł w tworzonym modelu, co utrudnia jego percepcję oraz interpretację. Okazuje się jednak, że bardzo duża liczba reguł w danym modelu, może wynikać zarówno z charakterystyki stosowanych algorytm-

mów uczenia maszynowego, jak i ze specyfiki analizowanej bazy przypadków [Preece i Shinghal, 1994; Spreeuwenberg i Gerrits, 2002; Ligęza, 2005].

Wspomniane wyżej cechy modeli uczenia oraz wspomniany paradygmat sztucznej inteligencji, stały się motywacją do podjęcia przeze mnie badań nad optymalizacją modeli w postaci zbioru reguł składniowych, uzyskanych poprzez zastosowanie *różnych* algorytmów uczenia maszynowego w odniesieniu do *tego samego* zbioru danych.

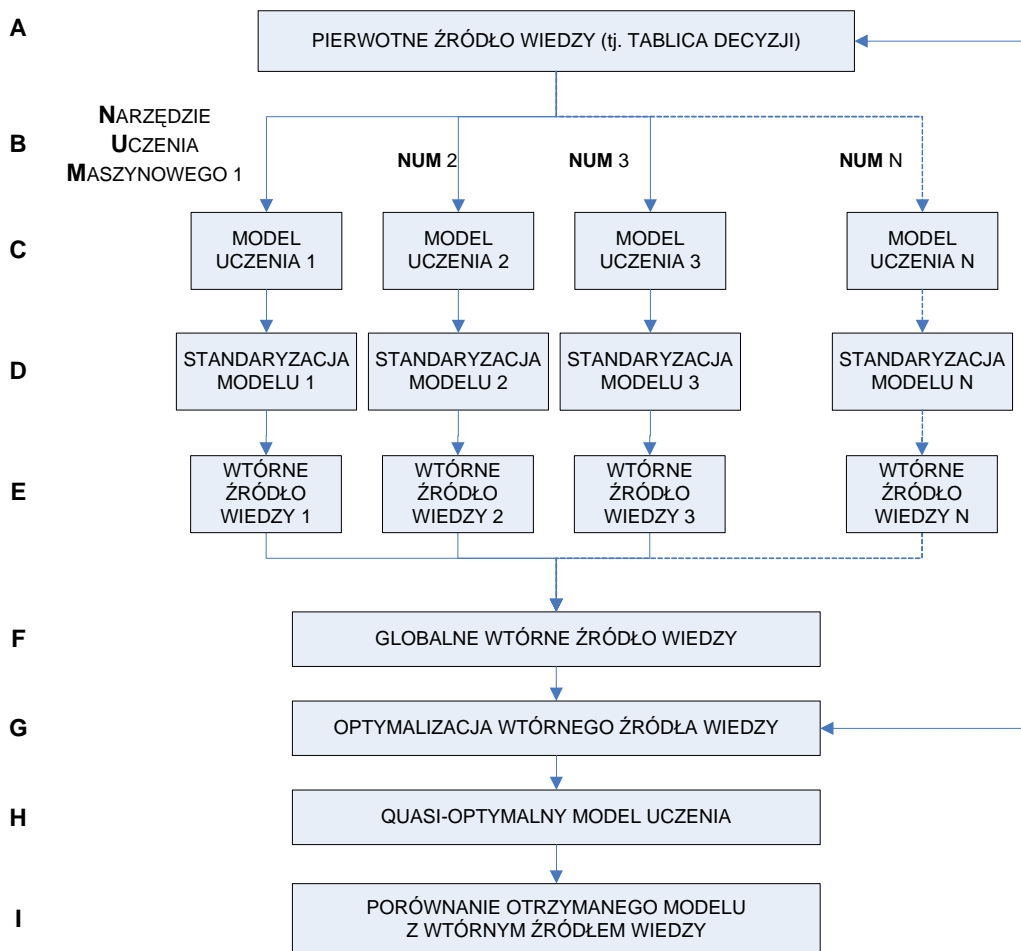
Badania przedstawione w niniejszej pracy stanowią – jak wspomniano – próbę opracowania metodologii optymalizacji różnych modeli uczenia, uzyskanych na podstawie tej samej tablicy decyzji [Hippe, Knap i Paja, 2002], (**A**, Rys. 5-1), nazwanej w tekście *pierwotnym źródłem wiedzy*. Modele tego typu, (**C**), powstają poprzez zastosowanie różnych algorytmów uczenia maszynowego, (**B**). Zakres badań obejmował będzie dwa odrębne podejścia: w pierwszym z nich przewiduje się **odrębną** optymalizację modeli uczenia, (**C**). Najpierw jednak, modele te muszą być poddane standaryzacji, (**D**), w sensie doprowadzenia do jednolitej formy zapisu reguł. Jako standard wybrano format reguł decyzji stosowany w algorytmie *GTS*, w którym reguły składniowe mają postać opisaną w Rozdziale 2, na str. 16. Otrzymany w ten sposób zestaw modeli uczenia tworzy tzw. *wtórne źródła wiedzy*, (**E**). Każde z nich zostanie poddane procesowi optymalizacji, (**F**), przy pomocy nowego, opracowanego w ramach rozprawy algorytmu. W efekcie przewiduje się otrzymanie zestawu quasi-optymalnych modeli uczenia, (**G**). Modele te mogą być wzajemnie porównane, a także można będzie obserwować skutki optymalizacji, (**F**), przez porównanie modeli **E** z modelami **G**. Porównanie to opierać się będzie na sprawdzeniu parametrów charakteryzujących wygenerowane modele uczenia; parametry te oraz proces optymalizacji wtórnych źródeł wiedzy zostaną szczegółowo przedstawione w dalszej części rozdziału.



Rys. 5-1 Idea zaproponowanej metody budowy optymalnych modeli uczenia na podstawie wtórnych źródeł wiedzy – podejście pierwsze

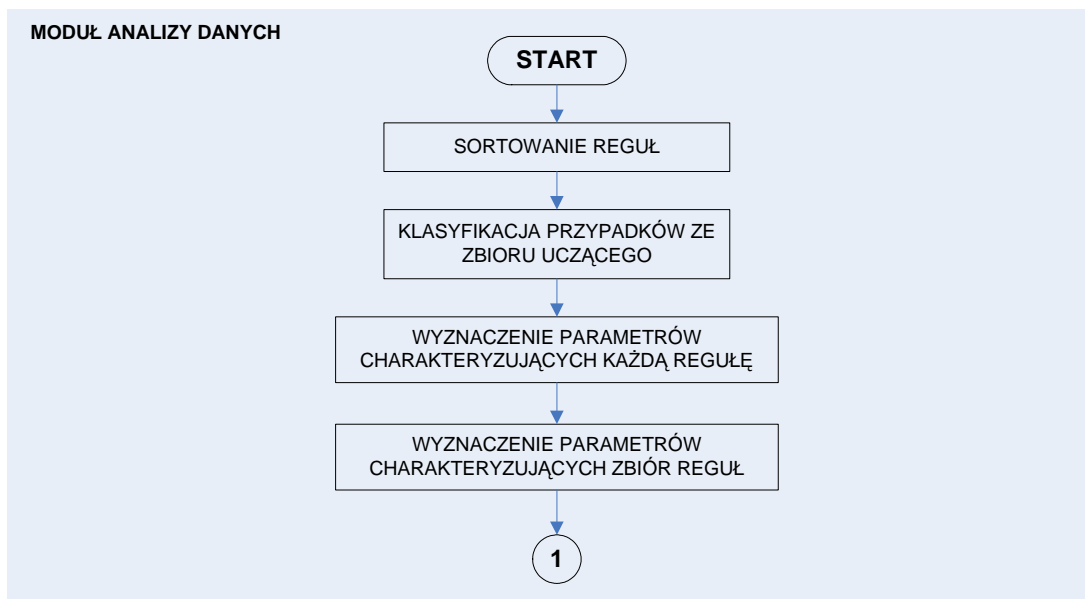
Natomiast w drugim podejściu (Rys. 5-2), uzyskane modele uczenia, **C**, po procesie standaryzacji, **D**, utworzą zestaw wtórnych źródeł wiedzy, **E**, które następnie zostaną połączone w globalne, wtórne źródło wiedzy, **(F)**. W następnym etapie, zostanie ono poddane procesowi optymalizacji, **(G)**, przy użyciu opracowanego algorytmu. W wyniku tych operacji zostanie utworzony globalny, quasi-optymalny model uczenia, **(H)**, aby następnie dokonać porównania, **(I)**, z poszczególnymi modelami składowymi, przy użyciu analogicznych parametrów jak w podejściu pierwszym. W obydwu opisanych ścieżkach, modele uczenia generowane są na podstawie tej samej tablicy decyzji **(A)**. Dlatego też w procesie optymalizacji, algorytm przetwarza nie tylko wtórne źródło wiedzy lecz odwołuje się również do pierwotnego źródła wiedzy (tj. tablicy decyzji), co uwidoczniło na obydwu sche-

matach. Natomiast zamysł optymalizacji modeli uczenia opiera się na koncepcji opracowania i zastosowania własnego algorytmu optymalizującego (G), wykonującego zestaw specyficznych operacji generycznych na zbiorze reguł decyzji.



Rys. 5-2 Idea zaproponowanej metody budowy optymalnych modeli uczenia na podstawie wtórnych źródeł wiedzy – podejście drugie

Ogólna postać opracowanego algorytmu optymalizacji wtórnych źródeł wiedzy składa się z trzech wzajemnie zależnych modułów: *analizy danych*, *optymalizacji* oraz *klasyfikacji*. Moduł analizy danych (Rys. 5-3) realizuje wstępne zadanie, polegające na przetworzeniu wczytanego pierwotnego źródła wiedzy (tablicy decyzji), symultanicznie z wczytaniem



Rys. 5-3 Algorytm optymalizacji wtórnych źródeł wiedzy – moduł analizy danych

i przetworzeniem odpowiedniego wtórnego źródła wiedzy, tj. zbioru reguł decyzji. Wczytane reguły są sortowane alfabetycznie według kategorii wskazywanych przez ich części konkluzyjne, według ilości warunków w regule oraz także sortowane są alfabetycznie warunki wewnątrz każdej reguły. Następnie dokonywana jest standardowa klasyfikacja przypadków z tablicy decyzji, polegająca na wykryciu przypadków, które są pokrywane przez analizowaną regułę. Na podstawie wyników tej klasyfikacji, zostaną wyznaczone parametry charakteryzujące każdą z reguł, a także ich średnie wartości opisujące cały zbiór reguł, a mianowicie: parametr *Siła*, który jest liczbą poprawnie klasyfikowanych przypadków z pierwotnego źródła wiedzy, *Specyficzność* – będący stosunkiem liczby poprawnie klasyfikowanych przypadków z bazy informacyjnej do liczby przypadków w tej bazie, należących do danej klasy decyzyjnej, parametr *Dokładność* – będący stosunkiem liczby poprawnie klasyfikowanych przypadków z bazy informacyjnej do sumy przypadków klasyfikowanych poprawnie oraz błędnie, i w końcu parametr *Ogólność*, który jest stosunkiem sumy przypadków klasyfikowanych poprawnie oraz błędnie do całkowitej liczby przypadków w zbiorze uczącym. Ostatni parametr to *Wsparcie*, czyli stosunek liczby poprawnie

klasyfikowanych przypadków z bazy informacyjnej do całkowitej liczby przypadków w tym zbiorze.

Kolejny moduł – moduł optymalizacji (Rys. 5-4) – obejmuje zestaw specyficznych operacji generycznych, które należy wykonać na zbiorze reguł decyzji w celu jego optymalizacji. Operacje te wiążą się z realizacją następujących zadań:

- *usunięcie reguł zbędnych* – tj. takich, które nie klasyfikują żadnego przypadku z bazy informacyjnej,
- *usunięcie reguł redundantnych* – tj. takich, które posiadają identyczne warunki i ich wartości w części warunkowej reguły, w ramach tej samej klasy decyzji,
- *usunięcie reguł pochłaniających* – tj. takich, które posiadają wspólną z innymi regułami część warunkową, wzbogaconą dodatkowo warunkami uzupełniającymi,
- *usunięcie warunków zbędnych* – tj. usunięcie warunków z danej reguły, które nie powodują zmiany liczby prawidłowo klasyfikowanych przypadków z bazy informacyjnej,
- *łączenie reguł* – tj. reguł zawierających ten sam zestaw atrybutów numerycznych w części warunkowej. Wartości tych atrybutów na ogół stanowią przedziały liczbowe, zawierające się w sobie lub zachodzące na siebie, natomiast zestaw atrybutów symbolicznych i ich wartości w części warunkowej musi być identyczny, przy czym cała operacja odbywa się w ramach tej samej klasy decyzji,
- *utworzenie reguł brakujących* – czasem obserwuje się, że opracowany model uczenia (zbiór reguł) nie klasyfikuje wszystkich przypadków z bazy informacyjnej. W celu uniknięcia tej sytuacji tworzone są tzw. reguły brakujące, zawierające warunki utworzone na podstawie występujących atry-

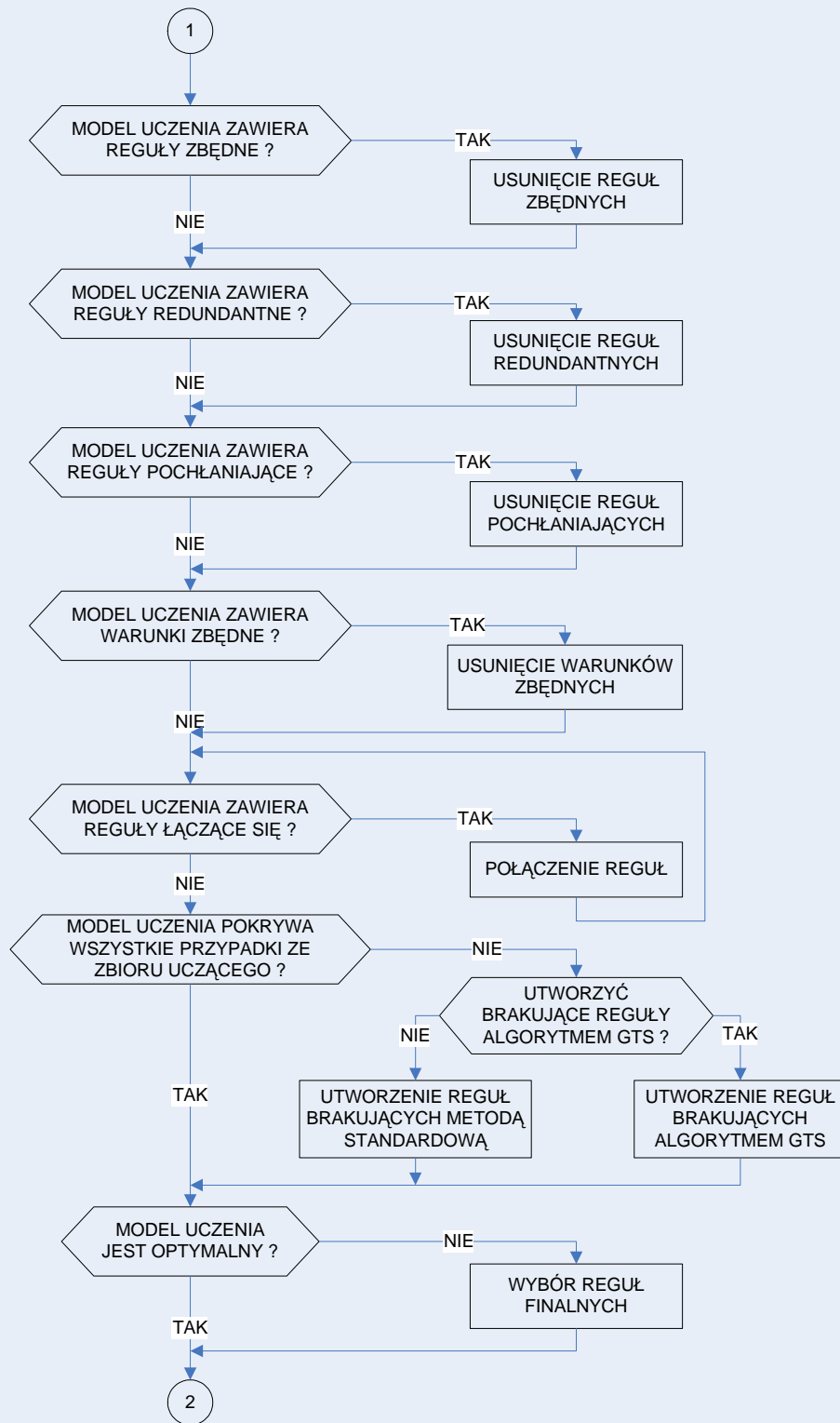
butów opisujących i ich wartości dla przypadków niesklasyfikowanych, lub alternatywnie, reguły te są generowane stosując algorytm pokrycia *GTS*,

- *wybór reguł finalnych* – będzie polegać na wyborze spośród zbioru wszystkich utworzonych, poprawnych reguł tylko tych, które posiadają największą *istotność H*. Istotność *H* reguły *r* obliczana jest następująco:

$$H(r) = [Siła(r) * Liczba warunków(r)] + Specyficzność(r) - Słabość(r) \quad (5-1)$$

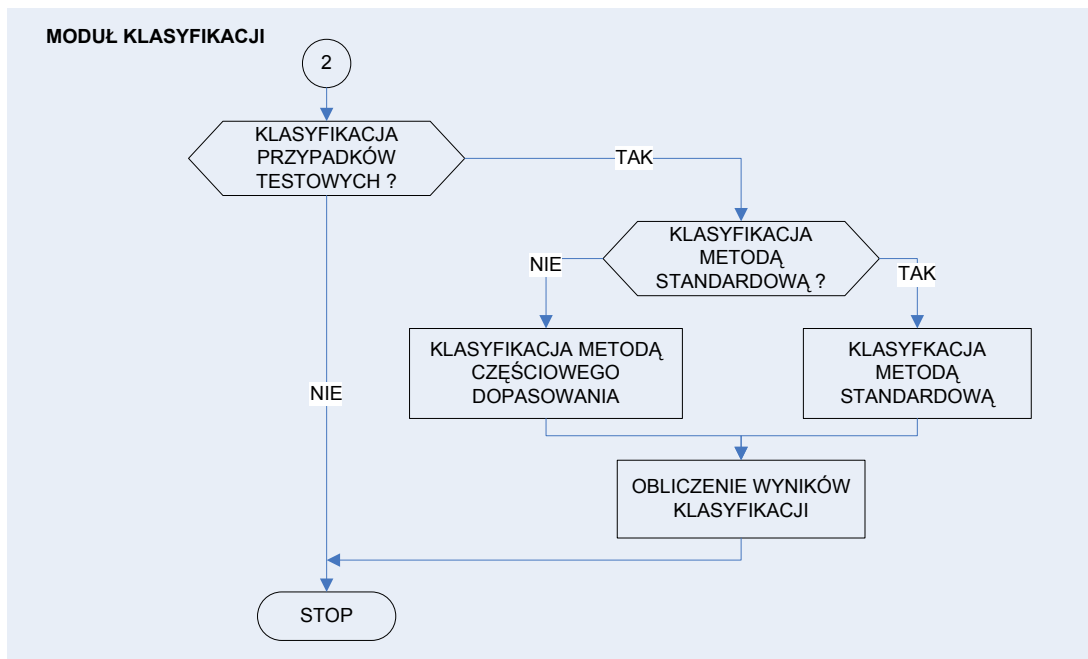
gdzie: *Liczba warunków(r)* jest liczbą warunków zawartych w części warunkowej reguły, *Słabość(r)* jest to stosunek liczby przypadków błędnie sklasyfikowanych do liczby przypadków sklasyfikowanych poprawnie. Pozostałe parametry były opisane wcześniej. Jednakże najważniejszym kryterium oceny modelu będzie *błąd klasyfikacji* (ang. *error rate*), określany jako stosunek liczby przykładów testowych błędnie sklasyfikowanych do liczby wszystkich przykładów testowych. Miara ta alternatywnie określana jest w procentach; im mniejsza wartość błędu, tym większa skuteczność klasyfikacyjna analizowanego zbioru reguł. Zakłada się przy tym, że przedstawione powyżej operacje przypuszczalnie doprowadzą do wyeliminowania wspomnianego w Rozdziale 4 zjawiska nadmiarowości w modelach uczenia, uzupełniając je nowymi, brakującymi regułami, a nawet spowodują utworzenie optymalnego (lub quasi-optymalnego) zbioru reguł decyzji.

MODUŁ OPTYMALIZACJI



Rys. 5-4 Algorytm optymalizacji wtórnych źródeł wiedzy – moduł optymalizacji

Uzyskane zoptymalizowane zbiory reguł zostaną zastosowane w kolejnym module tworzonego systemu (Rys. 5-5), do klasyfikacji przypadków testowych. Moduł ten wykorzystuje dwie metody klasyfikacji: pierwsza z nich, standardowa, polega na klasyfikacji obiektu nieznanego przy użyciu pierwszej reguły napotkanej w zbiorze reguł, która pokrywa atrybuty opisujące klasyfikowanego obiektu. Natomiast druga metoda, szczegółowo przedstawiona w Rozdziale 3, polega na tzw. *częściowym dopasowaniu* (ang. *partial matching*). Zoptymalizowany w ten sposób zbiór reguł wykorzystywany będzie do klasyfikacji testowego zbioru przypadków, w celu porównania charakteryzujących go parametrów oraz błędu klasyfikacji z odpowiednimi parametrami zbioru wejściowego (zbioru reguł przed optymalizacją). Porównanie wtórnych źródeł wiedzy z quasi-optymalnymi modelami uczenia opiera się na porównaniu charakteryzujących je parametrów, liczby przypadków niesklasyfikowanych, liczby przypadków sklasyfikowanych błędnie oraz błędu klasyfikacji. Moduł klasyfikacji pozwala również utworzyć tzw. *macierz rozproszenia*, ujawniającą rozkład prawidłowo i nieprawidłowo sklasyfikowanych przypadków w ramach danej klasy decyzji.



Rys. 5-5 Algorytm optymalizacji wtórnych źródeł wiedzy – moduł klasyfikacji

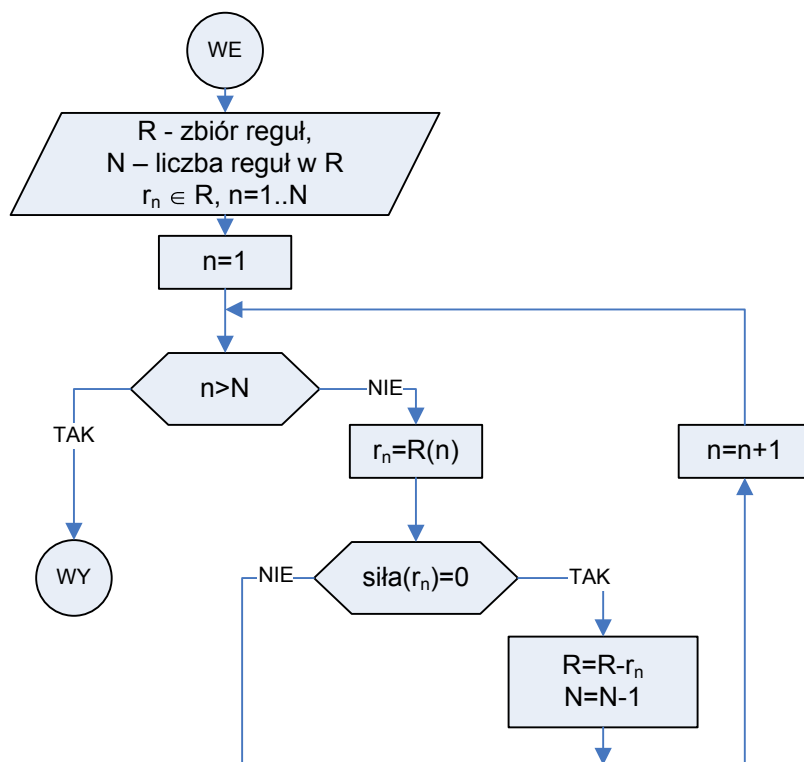
Do generowania reguł decyzji (tj. tworzenia wtórnych źródeł wiedzy) zastosowane będą trzy różne algorytmy uczenia maszynowego: **GTS** (General-To-Specific) [Hippe, 1997], **LEM2** [Grzymała-Busse, 1997] oraz **ID3/C4.5** [Quinlan, 1996; AI-TECH, 2000]. Natomiast jako pierwotne źródła wiedzy w prowadzonych badaniach zastosowane będą niektóre bazy informacyjne zaczerpnięte z repozytorium **UCI** [Newman i inni, 1998], oraz własna baza znamion melanocytowych skóry [Hippe i inni, 2003].

ROZDZIAŁ 6

GENERYCZNE OPERACJE OPTYMALIZACYJNE

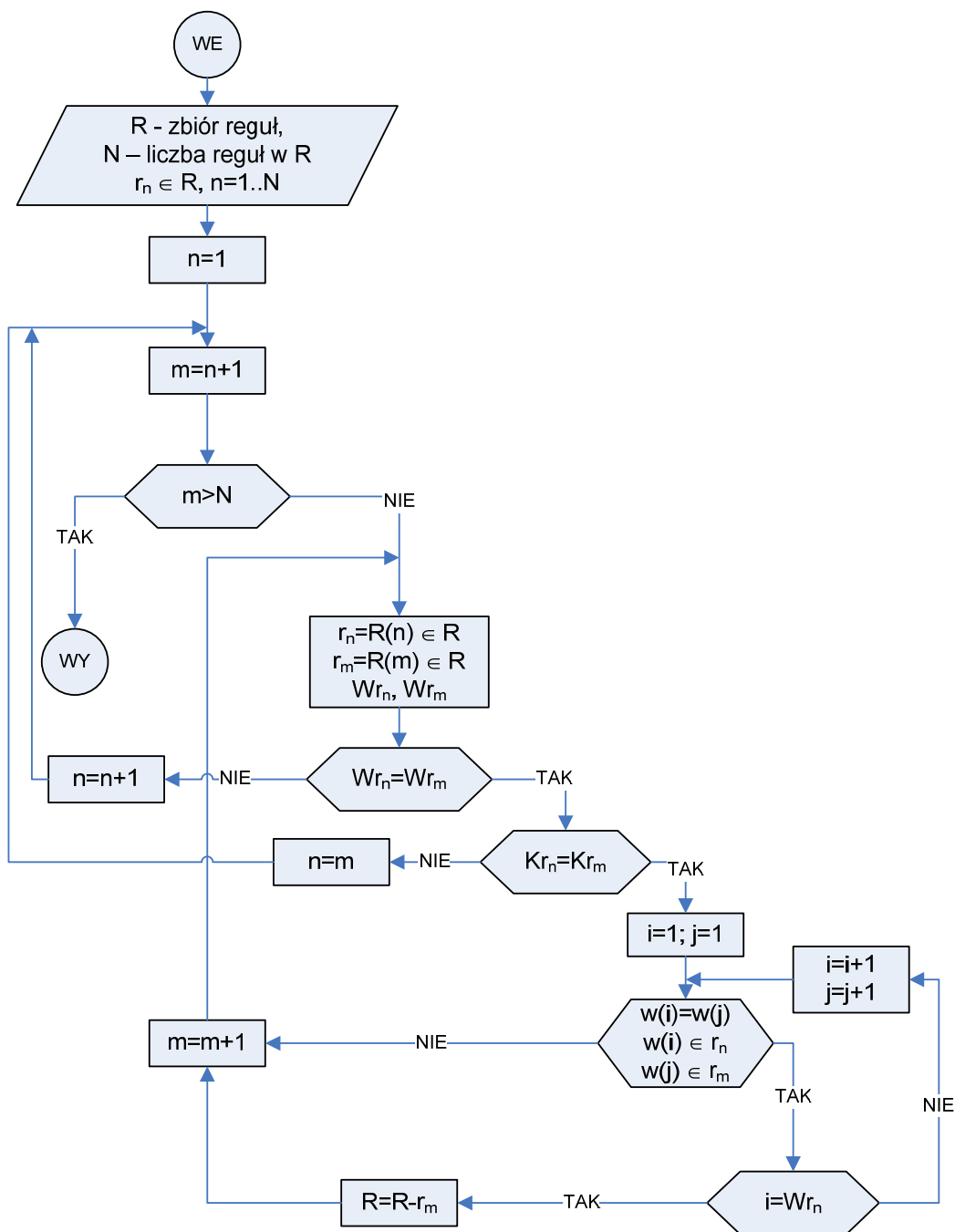
W poprzednim rozdziale przedstawiono koncepcję optymalizacji wtórnych źródeł wiedzy, w celu budowy optymalnych modeli uczenia. Koncepcja ta opiera się na zastosowaniu omówionych poprzednio procedur, z których najważniejszą jest sekwencyjne zastosowanie zestawu operacji generycznych, działających na zbiorze reguł R , utworzonym przez *moduł analizy danych* (Rozdział 5, Rys. 5-3).

Pierwszą z podstawowych operacji jakie można przeprowadzić na zbiorze reguł decyzji jest *usunięcie reguł zbędnych*, tj. takich, które nie klasyfikują żadnego przypadku z analizowanej bazy informacyjnej. Schemat algorytmu realizującego tę operację przedstawiono na Rys. 6-1. Każda z reguł decyzji r_n , należąca do zbioru reguł R , jest sprawdzana przy użyciu opisanego już parametru $Sila(r_n)$. W przypadku gdy parametr ten równy jest 0, reguła jest usuwana ze zbioru reguł R , tzn. $R=R-r_n$, przy czym iteracyjnie sprawdzane są kolejne reguły ze zbioru R .



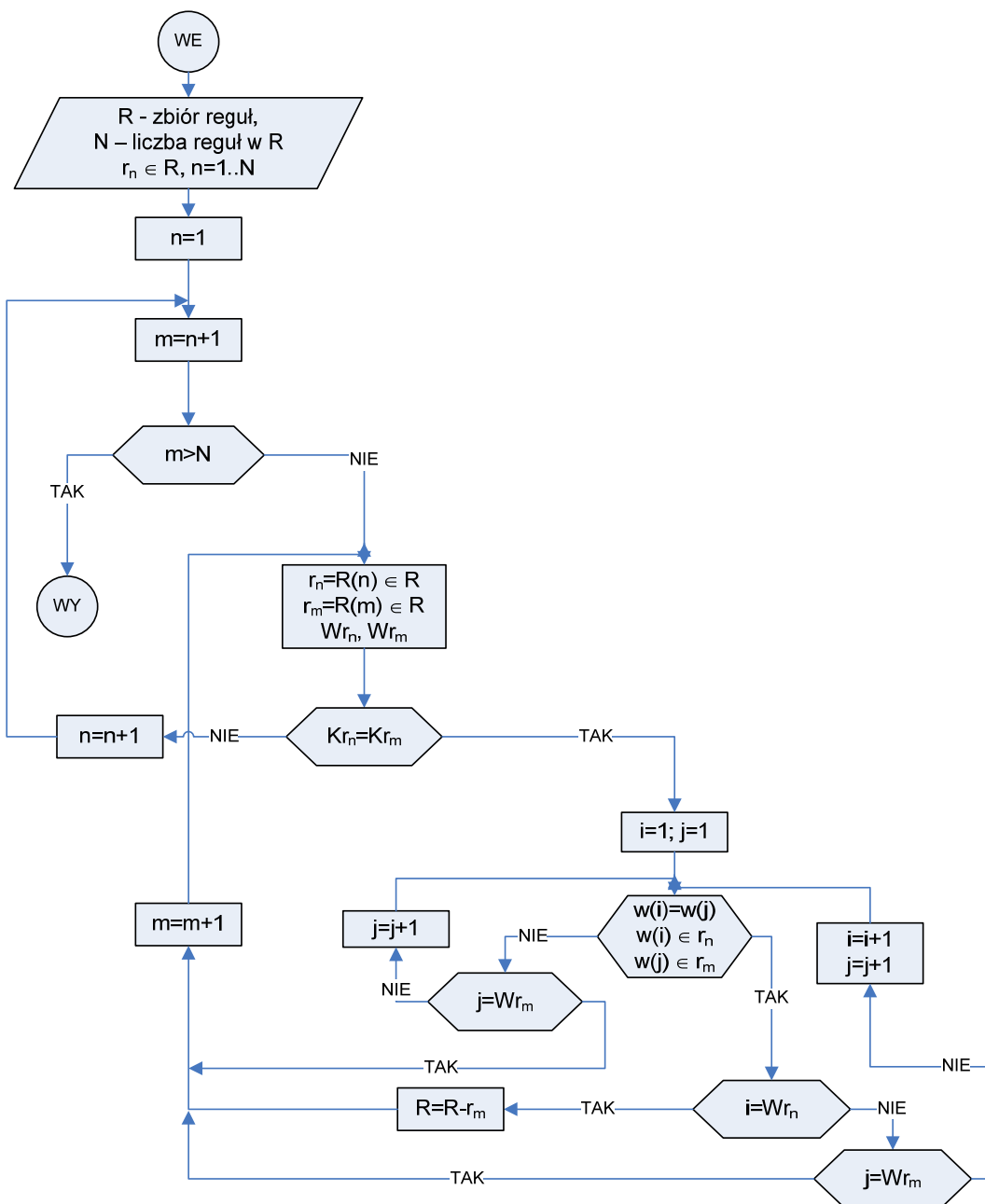
Rys. 6-1 Usunięcie reguł zbędnych – schemat algorytmu

Kolejną operacją jest *usunięcie reguł redundantnych*, tj. takich, które posiadają identyczne warunki w części warunkowej w obrębie tej samej klasy decyzji. Schemat działania tej procedury przedstawiono na Rys. 6-2. Każda reguła r_n , należąca do zbioru reguł R , jest porównywana z kolejnymi regułami r_m w obrębie tej samej kategorii przypadków, $Kr_n = Kr_m$, gdy liczba warunków obu reguł jest identyczna, $Wr_n = Wr_m$. Poszczególne warunki obydwu porównywanych reguł są wzajemnie porównywane, $w(i)=w(j)$, i gdy wszystkie warunki w pętli iteracyjnej spełniają tę zależność, oznacza to, że porównywane reguły są redundantne. Jedna z reguł, r_m , jest wtedy usuwana ze zbioru reguł R , tzn. $R=R-r_m$.



Rys. 6-2 Usunięcie reguł redundantnych – schemat algorytmu

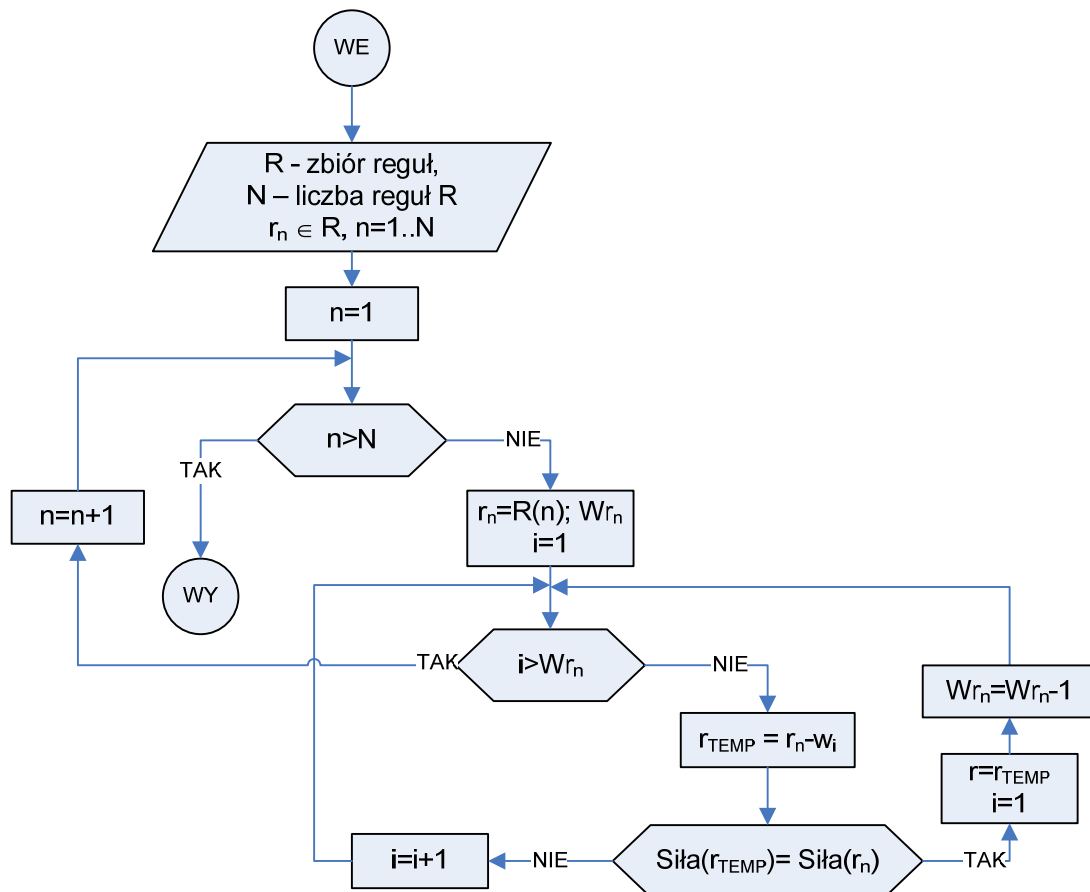
Podobną do omawianej operacją, jest *usunięcie reguł pochłaniających się* (Rys. 6-3), tj. takich, które posiadają wspólną z innymi regułami część warunkową, wzbogaconą dodatkowo warunkami uzupełniającymi.



Rys. 6-3 Usunięcie reguł pochłaniających się – schemat algorytmu

Operacja ta pozwala usunąć reguły, których część warunkowa zawiera w sobie część warunkową innej, porównywanej reguły, w ramach tej samej klasy decyzji, $Kr_n = Kr_m$. W porównaniu do algorytmu usuwania reguł redundantnych, liczba warunków w porównywanych regułach jest różna, $Wr_n \neq Wr_m$. Reguła bardziej szczegółowa, r_m , jest usuwana ze zbioru reguł R , tzn. $R=R-r_m$.

Wewnątrz każdej z reguł można natomiast zastosować operację nazwaną *usunięciem warunków zbędnych*, tj. takich, których usunięcie nie spowoduje zmiany liczby prawidłowo klasyfikowanych przypadków (Rys. 6-4).

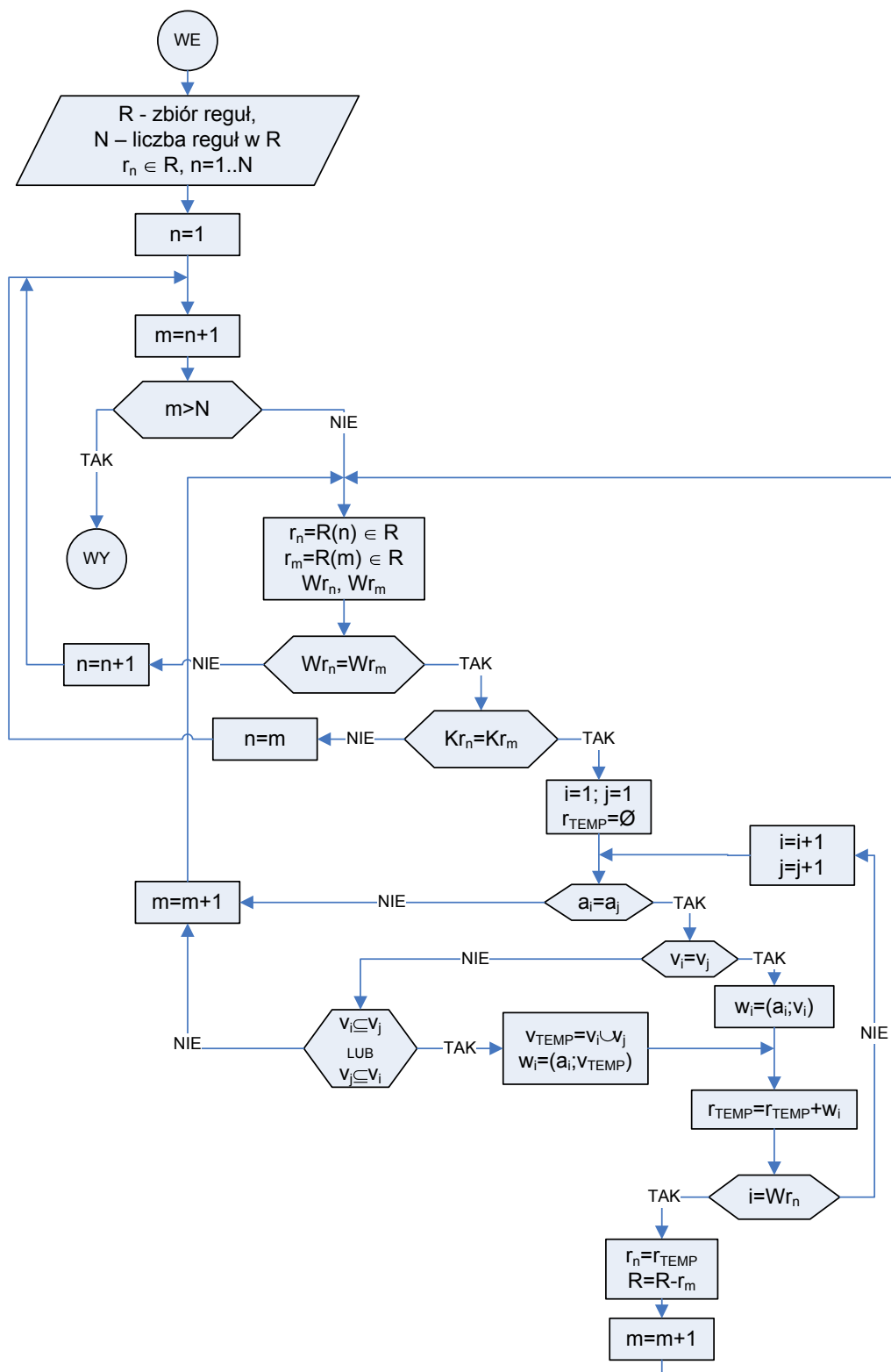


Rys. 6-4 Usunięcie warunków zbędnych – schemat algorytmu

Kolejne warunki w_i reguły r_n są odejmowane ze zbioru warunków składowych reguły; pozostałe warunki tworzą regułę tymczasową r_{TEMP} . Siła tej reguły jest porównywana z siłą reguły r_n ; gdy wartości te są równe, wówczas reguła r_{TEMP} zastępuje regułę r_n w zbiorze reguł R , tzn. $r = r_{TEMP}$.

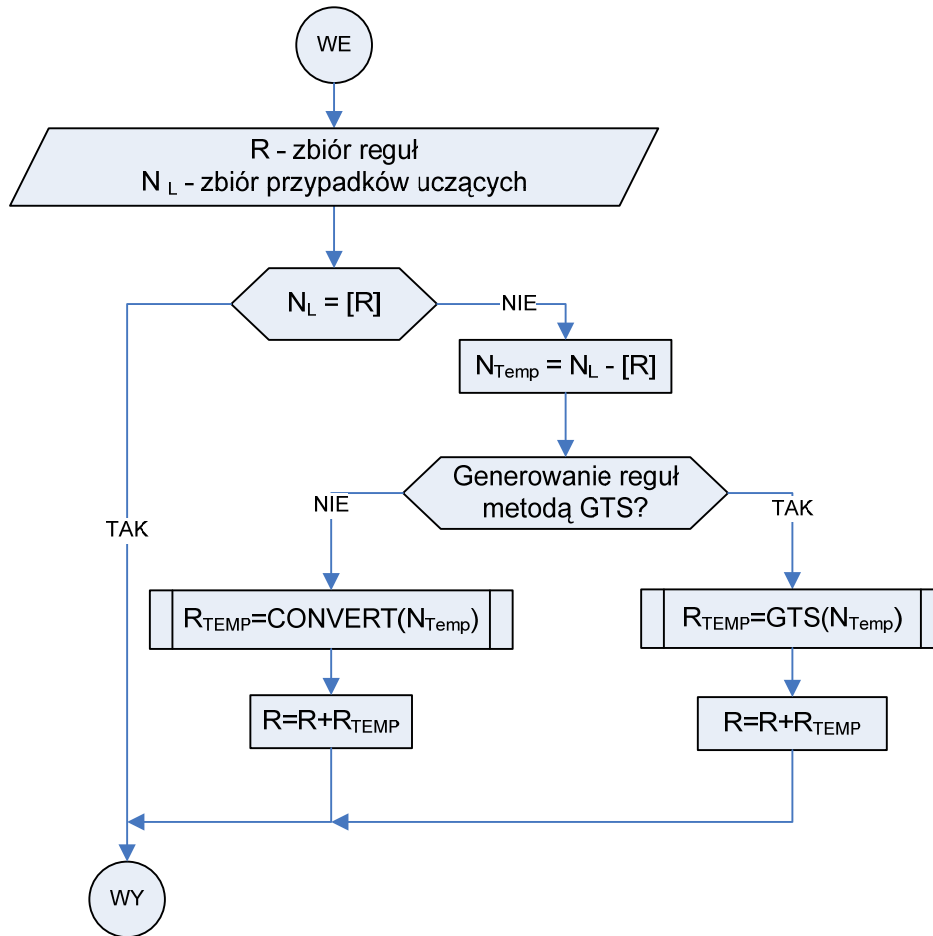
Kolejną operacją wykonywaną na zbiorze reguł, jest *łączenie reguł*, zawierających ten sam zestaw atrybutów numerycznych w części warunkowej (Rys. 6-5). Wartości tych atrybutów, v_i, v_j , na ogół stanowią pewne przedziały liczbowe zawierające się w sobie, lub zachodzące na siebie, natomiast zestaw atrybutów symbolicznych i ich wartości w części warunkowej musi być identyczny. Każda reguła r_n , należąca do zbioru reguł R , jest po-

równywana z innymi regułami r_m w obrębie tej samej klasy, $Kr_n = Kr_m$, gdy liczba warunków obydwu porównywanych reguł jest identyczna, $Wr_n = Wr_m$. W ramach obydwu rozpatrywanych reguł r_n , r_m , porównywane są kolejne warunki, a ściślej mówiąc zawarte w nich atrybuty a_i , a_j oraz ich wartości v_i , v_j . Jeżeli spełniony jest warunek, że $v_i \subseteq v_j$ lub $v_j \subseteq v_i$ wówczas tworzony jest warunek tymczasowy $v_{TEMP} = v_i \cup v_j$. Spełnienie tych warunków jest możliwe, gdy wartości logiczne porównywanych atrybutów tworzą przedziały numeryczne. W efekcie tej operacji, budowana jest reguła tymczasowa r_{TEMP} , zawierająca zsumowaną wartość połączonych atrybutów. Reguła ta zastępuje regułę r_n , natomiast reguła r_m jest usuwana ze zbioru reguł R .



Rys. 6-5 Łączenie reguł – schemat algorytmu

W zestawie operacji optymalizujących przewidziano również pewne operacje uzupełniające, mianowicie *utworzenie reguł brakujących* (Rys. 6-6).

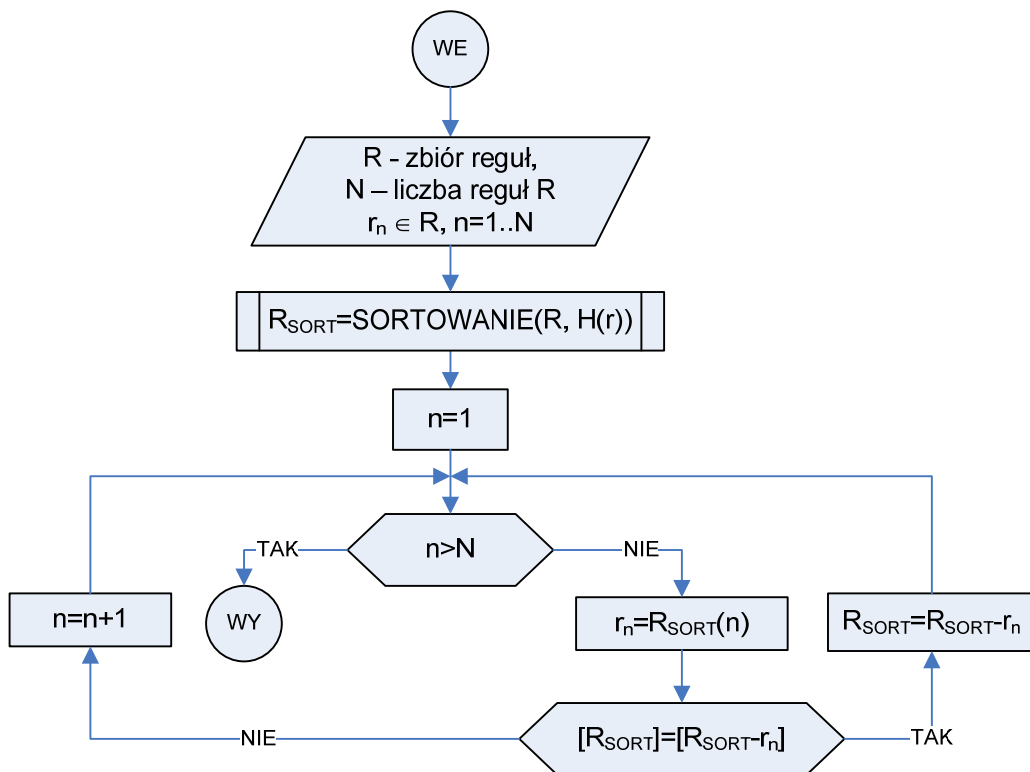


Rys. 6-6. Utworzenie reguł brakujących – schemat algorytmu

W tym celu, następuje sprawdzenie, czy zbiór przypadków N_L w bazie informacyjnej zawiera się w zbiorze $[R]$ przypadków pokrywanych przez zestaw reguł R . Jeżeli warunek ten nie jest spełniony, wówczas jest tworzony tymczasowy zbiór przypadków N_{TEMP} , będący różnicą $N_{TEMP} = N_L - [R]$. Zbiór ten zawiera zatem dane na podstawie których zostają utworzone reguły brakujące. Przyjęto założenie, że zostaną tu zastosowane dwie metody generowania reguł brakujących. W metodzie pierwszej, $CONVERT(N_{TEMP})$, przypadki ze zbioru N_{TEMP} są mapowane do postaci regułowej poprzez zamianę atrybutów i ich wartości, na odpowiednie warunki w regule. Metoda druga wykorzystuje algorytm pokrycia GTS , umożliwiającą utworzenie do-

datkowych reguł na podstawie zbioru przypadków, do tej pory niepokrytych, N_{TEMP} .

Końcową operacją jest *wybór reguł finalnych*, polegający na selekcji spośród zbioru wszystkich reguł R tych, które klasyfikują analizowane przypadki N_L , poprawnie rozpoznawane również przez inne reguły, posiadające mniejszą *istotność* H (Rozdział 5, wzór 5-1). Realizacja tej procedury wymaga, by zbiór reguł został uprzednio posortowany malejąco wg parametru H , $R_{SORT} = SORTOWANIE(R, H(r))$.



Rys. 6-7 Wybór reguł finalnych – schemat algorytmu

Teraz, w odniesieniu do każdej reguły r_n ze zbioru reguł R następuje sprawdzenie, czy zbiór przypadków pokrywanych przez zestaw reguł posortowanych $[R_{SORT}]$ jest taki sam jak zbiór przypadków pokrywanych przez zestaw reguł będących różnicą $[R_{SORT} - r_n]$. W przypadku spełnienia powyższego warunku, reguła r_n jest usuwana ze zbioru reguł, $R_{SORT} = R_{SORT} - r_n$. Ostatecznie uzyskiwany jest zbiór reguł finalnych.

Omówione w skrócie algorytmy optymalizacji zbioru reguł decyzji zostały zaimplementowane w opracowanym narzędziu informatycznym [Paja, 2005], ilustrującym główny cel niniejszej pracy. Wyniki pierwszych testów, przedstawione w następnym rozdziale, zdają się wskazywać, że opracowana w ramach niniejszej dysertacji metodologia optymalizacji regułowych modeli uczenia umożliwia uzyskanie całkowicie zadowalających rezultatów.

ROZDZIAŁ 7

BADANIA NAD OPTYMALIZACJĄ WTÓRNYCH ŹRÓDEŁ WIEDZY. DYSKUSJA WYNIKÓW

Opisane w niniejszym rozdziale badania miały na celu dokonanie obiektywnej weryfikacji i walidacji opracowanych algorytmów generowania quasi-optimalnych, wtórnych źródeł wiedzy. Zręby przyjętej metodyki tworzenia tych źródeł wiedzy przedstawiono w Rozdziałach 5 oraz 6. Badania zrealizowano via testowanie wybranych baz informacyjnych¹⁾, zaczerpniętych z repozytorium baz University of California, Irvine [Newman i inni, 1998], a także testowanie własnej bazy informacyjnej o znamionach melanocytowych skóry (baza o nazwie *P548134* [Hippe i inni, 2003], opisana w Załączniku A). Z uwagi na konieczność ograniczenia objętości rozprawy, w niniejszym rozdziale szczegółowo omówiono jedynie wyniki badań modeli uczenia utworzonych na podstawie analizy bazy *P548134* (symboliczno-numerycznej) oraz bazy *GLASS* (numerycznej). Przedstawione wyniki badań, a także sposób ich interpretacji zjawiskowej, można uznać za typowe, odnoszące się zatem również do wszystkich badanych baz informacyjnych. W ten sposób, jak się wydaje, można było uniknąć niewątpliwie bardzo uciążliwego śledzenia toku analizy wszystkich analizowanych baz informacyjnych oraz utworzonych na ich podstawie (stosując trzy odrębne

¹⁾ Badania były prowadzone poprzez analizę zbiorów danych, powszechnie używanych w badaniach z dziedziny uczenia maszynowego zawierających przypadki o wartościach numerycznych, symbolicznych lub mieszanych (numeryczno-symbolicznych). Były to zbiory: *Glass Identification*, *Car Evaluation Database*, *Heart Diseases*, *Wine* [Newman i inni, 1998].

algorytmy generowania reguł składniowych) wtórnych źródeł wiedzy. Taki sposób interpretacji uzyskanych wyników, umożliwił jednocześnie dokonanie interesujących spostrzeżeń na temat stosowanych algorytmów generowania modeli uczenia w postaci zbioru reguł typu JEŻELI ... TO.

Wyniki badań zebrano w Tablicach 7-1...7-8. Pierwsze cztery tablice (7-1...7-4) zawierają wyniki analizy bazy *P548134*, natomiast Tablice 7-5...7-8 odnoszą się do zbioru danych *GLASS*. Tablice 7-1...7-3 oraz 7-5...7-7 dotyczą metodologii badań, polegającej na odrębnej optymalizacji modeli uczenia, uzyskanych przy pomocy – jak wspomniano – różnych algorytmów uczenia maszynowego, a mianowicie algorytmu *GTS*, *LEM2* oraz *ID3*. Natomiast tablice 7-4 oraz 7-8 ujmują wyniki optymalizacji globalnego modelu uczenia, utworzonego przez połączenie modeli składowych *GTS+LEM2+ID3*. Każda ze wspomnianych tablic zawiera charakterystykę danego modelu uczenia oraz informacje o parametrach opisujących uzyskane zbiory reguł. Kolumna *Model początkowy* zawiera parametry modelu otrzymanego na podstawie analizy źródłowej bazy informacyjnej danego rodzaju. Kolumny *NU*, *RR*, *AR*, *W*, *L*, *NR* oraz *FR* zawierają parametry modelu, uzyskanego po zastosowaniu kolejnych operacji generycznych. Kolejność ich stosowania, zgodna z kolejnością wymienionych kolumn, została uznana za domyślną, i w tej kolejności były one wykonywane. Nie jest to jednak warunek konieczny, zatem operacje generyczne na regułach mogą być również wykonane w całkowicie dowolnej kolejności.

Analiza ilościowa badanych modeli uczenia

Na podstawie załączonych wyników badań można zauważyć, że różne algorytmy uczenia maszynowego generują odmienne modele uczenia. Modele te różnią się od siebie nie tylko licznością reguł, lecz także treścią ich ciała, co zostanie bardziej dokładnie przedstawione w dalszej części rozdziału, poświęconej analizie czynnika jakościowego utworzonych modeli uczenia.

Algorytm *GTS* w procesie analizy bazy informacyjnej *P548134*, buduje model zawierający 122 reguły, natomiast analiza tej samej bazy przez algorytm *LEM2* powoduje utworzenie zbioru 89 reguł decyzji. Z kolei, algorytm *ID3* generuje zbiór 67 reguł. Różnice liczby reguł w poszczególnych modelach uczenia są spowodowane nie tylko specyfiką działania zastosowanych algorytmów, lecz także metodą dyskretyzacji atrybutów numerycznych. Ponadto algorytm *LEM2* stosuje „wewnętrznie” pewne operacje optymalizujące, takie jak usunięcie reguł nadmiarowych oraz usunięcie nadmiarowych warunków. Modele uczenia, wygenerowane przez badane algorytmy, wykazują błąd klasyfikacji około 5.3%, 2.4% oraz 6.9%, odpowiednio.

Zastosowanie opisanych uprzednio operacji generycznych powoduje istotne zmiany charakterystyki modelu uczenia: ulega zmianie liczba reguł; tj. zmniejsza się dla algorytmu *GTS*, natomiast rośnie dla algorytmu *LEM2* i *ID3*. Równocześnie ulega zmianie błąd klasyfikacji w przypadku każdego z badanych algorytmów, ulegając zmniejszeniu do wartości 0%, 1.1% oraz 3.1%, odpowiednio. Porównując te wyniki z wynikami uzyskanymi w przypadku połączenia modeli uczenia w jedno, globalne źródło (Tablica 7-4) można zauważyć, że po optymalizacji modelu połączonego liczba reguł jest równa 103, natomiast błąd klasyfikacji otrzymanego modelu jest mniejszy lub porównywalny z wynikami poszczególnych modeli składowych, i wynosi tylko około 0.7%.

W odniesieniu do wyników badań drugiej, dyskutowanej tu bazy (*GLASS*), można je podsumować następująco: najwięcej reguł (156) generował algorytm *GTS*, model ten cechował się błędem klasyfikacji równym 15.9%. Najmniej reguł (tylko 31) wygenerował algorytm *ID3*, co ciekawe, błąd klasyfikacji (4.2%) był najmniejszy. Natomiast największy błąd klasyfikacji (34.1%), wykazał model otrzymany przy pomocy algorytmu *LEM2*: model ten zawierał 75 reguł decyzji. Po wykonaniu generycznych operacji optymalizujących liczba reguł w różnych modelach uległa następującym

zmianom: zmniejszyła się do 106 (algorytm *GTS*) oraz zwiększyła się do 80 (algorytm *LEM2*) i 39 (algorytm *ID3*). Natomiast dużym zmianom uległ błąd klasyfikacji, który odpowiednio zmniejszył się do wartości 0.47% (*GTS, LEM2*) oraz 3.74% (*ID3*). Można zatem uznać, że optymalizacja modelu uczenia spowodowała kilkunastokrotne zmniejszenie błędu klasyfikacji. Model globalny, utworzony poprzez połączenie poszczególnych modeli składowych *GTS+LEM2+ID3* (Tablica 7-8), po optymalizacji posiada porównywalne wartości parametrów, np. 85 reguł decyzji osiągających błąd klasyfikacji na poziomie 2.80%.

Tablica 7-1 Zmiany parametrów modelu uczenia po wykonaniu kolejnych operacji generycznej na przykładzie bazy P548134 – algorytm GTS.

| | | Wyniki po zastosowaniu kolejnej operacji generycznej | | | | | | | | | | |
|--------------------------------|--------------------------------------|--|------------------|-------------|-------------|-------------|-------------|-------------|-------------------|-------------|-------------|---------------|
| | | Model początkowy | I seria operacji | | | | | | II seria operacji | | | |
| | | | NU | RR | AR | W | Ł | NR | FR | W | Ł | FR |
| Charakterystyka modelu uczenia | liczba reguł | 122 | 107 | 107 | 107 | 107 | 105 | 136 | 119 | 119 | 118 | 113 |
| | ubytek warunków w zbiorze reguł | 0 | 0 | 0 | 0 | 40 | 0 | 0 | 0 | 108 | 0 | 0 |
| | liczba klasyfikowanych przypadków | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 |
| | liczba błędnych klasyfikacji | 29 | 29 | 29 | 29 | 29 | 29 | 0 | 0 | 0 | 0 | 0 |
| | liczba nieklasyfikowanych przypadków | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | błąd klasyfikacji | 5.3% | 5.3% | 5.3% | 5.3% | 5.3% | 5.3% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Parametry zbioru reguł | siła reguł | 9.148 | 10.43 | 10.43 | 10.43 | 10.626 | 10.63 | 9.456 | 9.429 | 10.101 | 10.13 | 10.221 |
| | dokładność reguł | 0.877 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | ogólność reguł | 0.017 | 0.019 | 0.019 | 0.019 | 0.019 | 0.019 | 0.017 | 0.017 | 0.018 | 0.018 | 0.019 |
| | specyficzność reguł | 0.065 | 0.074 | 0.074 | 0.074 | 0.075 | 0.075 | 0.067 | 0.069 | 0.074 | 0.074 | 0.075 |
| | wsparcie reguł | 0.017 | 0.019 | 0.019 | 0.019 | 0.019 | 0.019 | 0.017 | 0.017 | 0.018 | 0.018 | 0.019 |

gdzie:

kolumna **NU** – wyniki po zastosowaniu operacji usunięcia reguł zbędnych (nieużywanych);

kolumna **RR** – wyniki po zastosowaniu operacji usunięcia reguł redundantnych;

kolumna **AR** – wyniki po zastosowaniu operacji usunięcia reguł pochłaniających się;

kolumna **W** – wyniki po zastosowaniu operacji usunięcia zbędnych warunków;

kolumna **Ł** – wyniki po zastosowaniu operacji łączenia reguł;

kolumna **NR** – wyniki po zastosowaniu operacji dodania nowych reguł (brakujących);

kolumna **FR** – wyniki po zastosowaniu operacji wyboru reguł finalnych.

Tablica 7-2 Zmiany parametrów modelu uczenia po wykonaniu kolejnych operacji generycznej na przykładzie bazy **P548134** – algorytm **LEM2**.

| | | Model początkowy | Wyniki po zastosowaniu kolejnej operacji generycznej | | | | | | | |
|--------------------------------|--------------------------------------|------------------|--|-------------|-------------|-------------|-------------|-------------|-------------------|-------------|
| | | | I seria operacji | | | | | | II seria operacji | |
| | | | NU | RR | AR | W | Ł | NR | FR | W |
| Charakterystyka modelu uczenia | liczba reguł | 89 | 89 | 89 | 89 | 89 | 89 | 94 | 92 | 92 |
| | ubytek warunków w zbiorze reguł | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 9 |
| | liczba klasyfikowanych przypadków | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 |
| | liczba błędnych klasyfikacji | 13 | 13 | 13 | 13 | 13 | 13 | 1 | 6 | 6 |
| | liczba nieklasyfikowanych przypadków | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | błąd klasyfikacji | 2.4% | 2.4% | 2.4% | 2.4% | 2.4% | 2.4% | 0.2% | 1.1% | 1.1% |
| Parametry zbioru reguł | siła reguł | 11.708 | 11.708 | 11.708 | 11.708 | 11.708 | 11.708 | 11.617 | 11.707 | 11.88 |
| | dokładność reguł | 0.981 | 0.981 | 0.981 | 0.981 | 0.981 | 0.981 | 0.982 | 0.981 | 0.981 |
| | ogólność reguł | 0.022 | 0.022 | 0.022 | 0.022 | 0.022 | 0.022 | 0.022 | 0.022 | 0.022 |
| | specyficzność reguł | 0.083 | 0.083 | 0.083 | 0.083 | 0.083 | 0.083 | 0.084 | 0.084 | 0.086 |
| | wsparcie reguł | 0.021 | 0.021 | 0.021 | 0.021 | 0.021 | 0.021 | 0.021 | 0.021 | 0.022 |

gdzie:

kolumna **NU** – wyniki po zastosowaniu operacji usunięcia reguł zbędnych (nieużywanych);

kolumna **RR** – wyniki po zastosowaniu operacji usunięcia reguł redundantnych;

kolumna **AR** – wyniki po zastosowaniu operacji usunięcia reguł pochłaniających się;

kolumna **W** – wyniki po zastosowaniu operacji usunięcia zbędnych warunków;

kolumna **Ł** – wyniki po zastosowaniu operacji łączenia reguł;

kolumna **NR** – wyniki po zastosowaniu operacji dodania nowych reguł (brakujących);

kolumna **FR** – wyniki po zastosowaniu operacji wyboru reguł finalnych.

Tablica 7-3 Zmiany parametrów modelu uczenia po wykonaniu kolejnych operacji generycznej na przykładzie bazy **P548134** – algorytm **ID3**.

| | | Wyniki po zastosowaniu kolejnej operacji generycznej | | | | | | | | | | |
|--------------------------------|--------------------------------------|--|------------------|-------------|-------------|-------------|-------------|-------------|-------------------|-------------|-------------|---------------|
| | | Model początkowy | I seria operacji | | | | | | II seria operacji | | | |
| | | | NU | RR | AR | W | Ł | NR | FR | W | Ł | FR |
| Charakterystyka modelu uczenia | liczba reguł | 67 | 67 | 67 | 67 | 67 | 67 | 100 | 90 | 90 | 89 | 85 |
| | ubytek warunków w zbiorze reguł | 0 | 0 | 0 | 0 | 159 | 0 | 0 | 0 | 128 | 0 | 0 |
| | liczba klasyfikowanych przypadków | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 |
| | liczba błędnych klasyfikacji | 38 | 38 | 38 | 38 | 35 | 35 | 22 | 20 | 21 | 21 | 17 |
| | liczba nieklasyfikowanych przypadków | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | błąd klasyfikacji | 6.9% | 6.9% | 6.9% | 6.9% | 6.4% | 6.4% | 4.0% | 3.7% | 3.8% | 3.8% | 3.1% |
| Parametry zbioru reguł | siła reguł | 7.612 | 7.612 | 7.612 | 7.612 | 15.254 | 15.254 | 12.08 | 11.444 | 12.378 | 12.47 | 12.624 |
| | dokładność reguł | 0.887 | 0.887 | 0.887 | 0.887 | 0.937 | 0.937 | 0.958 | 0.958 | 0.958 | 0.958 | 0.961 |
| | ogólność reguł | 0.015 | 0.015 | 0.015 | 0.015 | 0.029 | 0.029 | 0.023 | 0.022 | 0.023 | 0.023 | 0.024 |
| | specyficzność reguł | 0.055 | 0.055 | 0.055 | 0.055 | 0.106 | 0.106 | 0.086 | 0.083 | 0.089 | 0.09 | 0.09 |
| | wsparcie reguł | 0.014 | 0.014 | 0.014 | 0.014 | 0.028 | 0.028 | 0.022 | 0.021 | 0.023 | 0.023 | 0.023 |

gdzie:

kolumna **NU** – wyniki po zastosowaniu operacji usunięcia reguł zbędnych (nieużywanych);

kolumna **RR** – wyniki po zastosowaniu operacji usunięcia reguł redundantnych;

kolumna **AR** – wyniki po zastosowaniu operacji usunięcia reguł pochłaniających się;

kolumna **W** – wyniki po zastosowaniu operacji usunięcia zbędnych warunków;

kolumna **Ł** – wyniki po zastosowaniu operacji łączenia reguł;

kolumna **NR** – wyniki po zastosowaniu operacji dodania nowych reguł (brakujących);

kolumna **FR** – wyniki po zastosowaniu operacji wyboru reguł finalnych.

Tablica 7-4 Zmiany parametrów modelu uczenia po wykonaniu kolejnych operacji generycznej na przykładzie bazy P548134 – model globalny (GTS+LEM2+ID3).

| | | Model początkowy | Wyniki po zastosowaniu kolejnej operacji generycznej | | | | | | | |
|--------------------------------|--------------------------------------|------------------|--|-------------|-------------|-------------|-------------|-------------|-------------------|-------------|
| | | | I seria operacji | | | | | | II seria operacji | |
| | | | NU | RR | AR | W | Ł | NR | FR | W |
| Charakterystyka modelu uczenia | liczba reguł | 278 | 263 | 252 | 242 | 242 | 235 | 236 | 103 | 103 |
| | ubytek warunków w zbiorze reguł | 0 | 0 | 0 | 0 | 178 | 0 | 0 | 0 | 4 |
| | liczba klasyfikowanych przypadków | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 | 548 |
| | liczba błędnych klasyfikacji | 13 | 13 | 13 | 13 | 16 | 16 | 16 | 4 | 4 |
| | liczba nieklasyfikowanych przypadków | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | błąd klasyfikacji | 2.4% | 2.4% | 2.4% | 2.4% | 2.9% | 2.9% | 2.9% | 0.7% | 0.7% |
| Parametry zbioru reguł | siła reguł | 9.597 | 10.144 | 9.357 | 9.207 | 11.112 | 10.979 | 10.941 | 8.951 | 9.00 |
| | dokładność reguł | 0.913 | 0.965 | 0.963 | 0.962 | 0.976 | 0.975 | 0.975 | 0.985 | 0.985 |
| | ogólność reguł | 0.018 | 0.019 | 0.018 | 0.017 | 0.021 | 0.02 | 0.02 | 0.017 | 0.017 |
| | specyficzność reguł | 0.068 | 0.072 | 0.066 | 0.065 | 0.078 | 0.077 | 0.077 | 0.066 | 0.067 |
| | wsparcie reguł | 0.018 | 0.019 | 0.017 | 0.017 | 0.02 | 0.02 | 0.02 | 0.016 | 0.016 |

gdzie:

kolumna **NU** – wyniki po zastosowaniu operacji usunięcia reguł zbędnych (nieużywanych);

kolumna **RR** – wyniki po zastosowaniu operacji usunięcia reguł redundantnych;

kolumna **AR** – wyniki po zastosowaniu operacji usunięcia reguł pochłaniających się;

kolumna **W** – wyniki po zastosowaniu operacji usunięcia zbędnych warunków;

kolumna **Ł** – wyniki po zastosowaniu operacji łączenia reguł;

kolumna **NR** – wyniki po zastosowaniu operacji dodania nowych reguł (brakujących);

kolumna **FR** – wyniki po zastosowaniu operacji wyboru reguł finalnych.

Tablica 7-5 Zmiany parametrów modelu uczenia po wykonaniu kolejnych operacji generycznej na przykładzie bazy **GLASS** – algorytm **GTS**.

| | | Model początkowy | Wyniki po zastosowaniu kolejnej operacji generycznej | | | | | | | | |
|--------------------------------|--------------------------------------|------------------|--|--------------|--------------|--------------|--------------|-------------|-------------------|-------------|--------------|
| | | | I seria operacji | | | | | | II seria operacji | | |
| | | | NU | RR | AR | W | Ł | NR | FR | W | FR |
| Charakterystyka modelu uczenia | liczba reguł | 156 | 129 | 129 | 129 | 129 | 118 | 143 | 108 | 108 | 106 |
| | ubytek warunków w zbiorze reguł | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 97 | 0 |
| | liczba klasyfikowanych przypadków | 190 | 190 | 190 | 190 | 190 | 190 | 214 | 214 | 214 | 214 |
| | liczba błędnych klasyfikacji | 10 | 10 | 10 | 10 | 10 | 10 | 4 | 1 | 1 | 1 |
| | liczba nieklasyfikowanych przypadków | 24 | 24 | 24 | 24 | 24 | 24 | 0 | 0 | 0 | 0 |
| | błąd klasyfikacji | 15.9% | 15.9% | 15.9% | 15.9% | 15.9% | 15.9% | 1.9% | 0.5% | 0.5% | 0.5% |
| Parametry zbioru reguł | siła reguł | 1.353 | 1.636 | 1.636 | 1.636 | 1.636 | 1.788 | 2.189 | 2.278 | 2.315 | 2.33 |
| | dokładność reguł | 0.796 | 0.963 | 0.963 | 0.963 | 0.963 | 0.959 | 0.967 | 0.971 | 0.971 | 0.971 |
| | ogólność reguł | 0.007 | 0.008 | 0.008 | 0.008 | 0.008 | 0.009 | 0.011 | 0.011 | 0.011 | 0.011 |
| | specyficzność reguł | 0.035 | 0.043 | 0.043 | 0.043 | 0.043 | 0.047 | 0.059 | 0.061 | 0.062 | 0.062 |
| | wsparcie reguł | 0.006 | 0.008 | 0.008 | 0.008 | 0.008 | 0.008 | 0.01 | 0.011 | 0.011 | 0.011 |

gdzie:

kolumna **NU** – wyniki po zastosowaniu operacji usunięcia reguł zbędnych (nieużywanych);

kolumna **RR** – wyniki po zastosowaniu operacji usunięcia reguł redundantnych;

kolumna **AR** – wyniki po zastosowaniu operacji usunięcia reguł pochłaniających się;

kolumna **W** – wyniki po zastosowaniu operacji usunięcia zbędnych warunków;

kolumna **Ł** – wyniki po zastosowaniu operacji łączenia reguł;

kolumna **NR** – wyniki po zastosowaniu operacji dodania nowych reguł (brakujących);

kolumna **FR** – wyniki po zastosowaniu operacji wyboru reguł finalnych.

Tablica 7-6 Zmiany parametrów modelu uczenia po wykonaniu kolejnych operacji generycznej na przykładzie bazy **GLASS** – algorytm **LEM2**.

| | Model początkowy | Wyniki po zastosowaniu kolejnej operacji generycznej | | | | | | | | | | |
|--------------------------------|--------------------------------------|--|--------------|--------------|--------------|--------------|--------------|-------------|-------------------|-------------|-------------|--------------|
| | | I seria operacji | | | | | | | II seria operacji | | | |
| | | NU | RR | AR | W | Ł | NR | FR | W | Ł | FR | |
| Charakterystyka modelu uczenia | liczba reguł | 75 | 42 | 42 | 42 | 42 | 41 | 102 | 83 | 83 | 82 | 80 |
| | ubytek warunków w zbiorze reguł | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 180 | 0 | 0 |
| | liczba klasyfikowanych przypadków | 214 | 214 | 214 | 214 | 214 | 214 | 214 | 214 | 214 | 214 | 214 |
| | liczba błędnych klasyfikacji | 73 | 73 | 73 | 73 | 73 | 79 | 2 | 1 | 1 | 1 | 1 |
| | liczba nieklasyfikowanych przypadków | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | błąd klasyfikacji | 34.1% | 34.1% | 34.1% | 34.1% | 34.1% | 36.9% | 0.9% | 0.5% | 0.5% | 0.5% | 0.5% |
| Parametry zbioru reguł | siła reguł | 1.72 | 3.071 | 3.071 | 3.071 | 3.071 | 3.049 | 3.382 | 3.193 | 3.313 | 3.354 | 3.237 |
| | dokładność reguł | 0.557 | 0.994 | 0.994 | 0.994 | 0.994 | 0.99 | 0.996 | 0.997 | 0.997 | 0.997 | 0.997 |
| | ogólność reguł | 0.008 | 0.014 | 0.014 | 0.014 | 0.014 | 0.014 | 0.016 | 0.015 | 0.016 | 0.016 | 0.015 |
| | specyficzność reguł | 0.037 | 0.066 | 0.066 | 0.066 | 0.066 | 0.067 | 0.092 | 0.085 | 0.088 | 0.089 | 0.088 |
| | wsparcie reguł | 0.008 | 0.014 | 0.014 | 0.014 | 0.014 | 0.014 | 0.016 | 0.015 | 0.015 | 0.016 | 0.015 |

gdzie:

kolumna **NU** – wyniki po zastosowaniu operacji usunięcia reguł zbędnych (nieużywanych);

kolumna **RR** – wyniki po zastosowaniu operacji usunięcia reguł redundantnych;

kolumna **AR** – wyniki po zastosowaniu operacji usunięcia reguł pochłaniających się;

kolumna **W** – wyniki po zastosowaniu operacji usunięcia zbędnych warunków;

kolumna **Ł** – wyniki po zastosowaniu operacji łączenia reguł;

kolumna **NR** – wyniki po zastosowaniu operacji dodania nowych reguł (brakujących);

kolumna **FR** – wyniki po zastosowaniu operacji wyboru reguł finalnych.

Tablica 7-7 Zmiany parametrów modelu uczenia po wykonaniu kolejnych operacji generycznej na przykładzie bazy **GLASS** – algorytm **ID3**.

| | | Model początkowy | Wyniki po zastosowaniu kolejnej operacji generycznej | | | | | | | |
|--------------------------------|--------------------------------------|------------------|--|-------------|-------------|-------------|-------------|-------------|-------------|-------------------|
| | | | I seria operacji | | | | | | | II seria operacji |
| | | | NU | RR | AR | W | Ł | NR | FR | W |
| Charakterystyka modelu uczenia | liczba reguł | 31 | 31 | 31 | 31 | 31 | 31 | 40 | 39 | 39 |
| | ubytek warunków w zbiorze reguł | 0 | 0 | 0 | 0 | 46 | 0 | 0 | 0 | 30 |
| | liczba klasyfikowanych przypadków | 214 | 214 | 214 | 214 | 214 | 214 | 214 | 214 | 214 |
| | liczba błędnych klasyfikacji | 9 | 9 | 9 | 9 | 9 | 9 | 6 | 6 | 8 |
| | liczba nieklasyfikowanych przypadków | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | błąd klasyfikacji | 4.2% | 4.2% | 4.2% | 4.2% | 4.2% | 4.2% | 2.8% | 2.8% | 3.7% |
| Parametry zbioru reguł | siła reguł | 6.613 | 6.613 | 6.613 | 6.613 | 7.032 | 7.032 | 5.775 | 5.872 | 5.949 |
| | dokładność reguł | 0.936 | 0.936 | 0.936 | 0.936 | 0.942 | 0.942 | 0.955 | 0.962 | 0.962 |
| | ogólność reguł | 0.032 | 0.032 | 0.032 | 0.032 | 0.034 | 0.034 | 0.028 | 0.028 | 0.029 |
| | specyficzność reguł | 0.183 | 0.183 | 0.183 | 0.183 | 0.191 | 0.191 | 0.161 | 0.165 | 0.167 |
| | wsparcie reguł | 0.031 | 0.031 | 0.031 | 0.031 | 0.033 | 0.033 | 0.027 | 0.027 | 0.028 |

gdzie:

kolumna **NU** – wyniki po zastosowaniu operacji usunięcia reguł zbędnych (nieużywanych);

kolumna **RR** – wyniki po zastosowaniu operacji usunięcia reguł redundantnych;

kolumna **AR** – wyniki po zastosowaniu operacji usunięcia reguł pochłaniających się;

kolumna **W** – wyniki po zastosowaniu operacji usunięcia zbędnych warunków;

kolumna **Ł** – wyniki po zastosowaniu operacji łączenia reguł;

kolumna **NR** – wyniki po zastosowaniu operacji dodania nowych reguł (brakujących);

kolumna **FR** – wyniki po zastosowaniu operacji wyboru reguł finalnych.

Tablica 7-8 Zmiany parametrów modelu uczenia po wykonaniu kolejnych operacji generycznej na przykładzie bazy **GLASS** – model globalny (**GTS+LEM2+ID3**).

| | | Wyniki po zastosowaniu kolejnej operacji generycznej | | | | | | | |
|--------------------------------|--------------------------------------|--|------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | | Model początkowy | I seria operacji | | | | | | |
| | | | NU | RR | AR | W | Ł | NR | FR |
| Charakterystyka modelu uczenia | liczba reguł | 262 | 201 | 201 | 201 | 201 | 189 | 189 | 85 |
| | ubytek warunków w zbiorze reguł | 0 | 0 | 0 | 0 | 50 | 0 | 0 | 0 |
| | liczba klasyfikowanych przypadków | 214 | 214 | 214 | 214 | 214 | 214 | 214 | 214 |
| | liczba błędnych klasyfikacji | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 6 |
| | liczba nieklasyfikowanych przypadków | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | błąd klasyfikacji | 3.3% | 3.3% | 3.3% | 3.3% | 3.3% | 3.3% | 3.3% | 2.8% |
| Parametry zbioru reguł | siła reguł | 2.076 | 2.706 | 2.706 | 2.706 | 2.801 | 2.947 | 2.947 | 3 |
| | dokładność reguł | 0.737 | 0.961 | 0.961 | 0.961 | 0.962 | 0.959 | 0.959 | 0.957 |
| | ogólność reguł | 0.01 | 0.013 | 0.013 | 0.013 | 0.014 | 0.015 | 0.015 | 0.015 |
| | specyficzność reguł | 0.053 | 0.069 | 0.069 | 0.069 | 0.071 | 0.075 | 0.075 | 0.078 |
| | wsparcie reguł | 0.01 | 0.013 | 0.013 | 0.013 | 0.013 | 0.014 | 0.014 | 0.014 |

gdzie:

kolumna **NU** – wyniki po zastosowaniu operacji usunięcia reguł zbędnych (nieużywanych);

kolumna **RR** – wyniki po zastosowaniu operacji usunięcia reguł redundantnych;

kolumna **AR** – wyniki po zastosowaniu operacji usunięcia reguł pochłaniających się;

kolumna **W** – wyniki po zastosowaniu operacji usunięcia zbędnych warunków;

kolumna **Ł** – wyniki po zastosowaniu operacji łączenia reguł;

kolumna **NR** – wyniki po zastosowaniu operacji dodania nowych reguł (brakujących);

kolumna **FR** – wyniki po zastosowaniu operacji wyboru reguł finalnych.

Analiza jakościowa modelu uczenia „znamiona melanocytowe skóry”

Przeprowadzona do tej pory dyskusja uzyskanych wyników badań koncentrowała się głównie na zagadnieniach ilościowych, konfrontując liczby reguł oraz błąd klasyfikacji badanych modeli. Dodatkowe informacje ilościowe, nie wymagające odrębnego komentarza, zawiera każda z załączonych tablic, w postaci parametrów oceny zbioru reguł w kolejnych, stopniowo optymalizowanych modelach, a mianowicie siły, dokładności, ogólności, specyficzności oraz wsparcia reguł.

Jakościową analizę reguł przed i po zastosowaniu operacji generycznych przeprowadzono na przykładzie modelu uczenia generowanego algorytmem *GTS*, w odniesieniu do bazy informacyjnej znamion melanocytowych skóry (*P548134*). Ograniczenie procesu jakościowej analizy do omówienia jedynie tego tylko przykładu jest z zamysłem spowodowane dwoma przyczynami:

(i) model dotyczący znamion melanocytowych skóry – jak pokazały przeprowadzone badania – okazał się modelem najbardziej skomplikowanym, zatem stanowi dobre tło do objaśnienia metodyki analizy i oceny skutków optymalizacji poprzez zastosowanie zbioru operacji generycznych. Zbiory reguł innych badanych ciągów uczących analizuje się analogicznie,

(ii) elementy tego właśnie modelu uczenia, związanego integralnie z głównym nurtem badań prowadzonych w Katedrze Systemów Ekspertowych i Sztucznej Inteligencji w Wyższej Szkole Informatyki i Zarządzania w Rzeszowie, zostały z powodzeniem zaimplementowane w systemie *IMDLS* (*Internet Melanoma Diagnosing and Learning System*) [Hippe i inni, 2005a; Hippe i inni, 2005b], dostępnym w sieci internetowej pod adresem <http://www.melanoma.pl> uzyskując w ten sposób możliwość ciągłej obserwacji działania opracowanego modelu.

W Tablicy 7-9 oraz 7-10 przedstawiono informację na temat liczby reguł, zawierających określoną liczbę warunków (od 1 do 11) w odniesieniu do identyfikowanych kategorii znamion melanocytowych: *znamię łagodne*, *znamię błękitne*, *znamię podejrzane* oraz *znamię złośliwe*. Tablica 7-9 zawiera informacje dotyczące modelu uczenia przed optymalizacją, natomiast Tablica 7-10 ujmuje te same informacje po optymalizacji. Analizując wyniki badań zawarte w wymienionych tablicach można zauważyć, że proces optymalizacji spowodował usunięcie większości bardzo szczegółowych reguł zawierających 9, 10 lub 11 warunków. Inaczej mówiąc, nastąpiło pewne uogólnienie modelu uczenia, reguły stały się bardziej ogólne, przy jednoczesnym zmniejszeniu ich liczby (ze 122 do 113) oraz zmniejszeniu błędu klasyfikacji z wartości około 5.3% do 0% (Tablica 7-1).

Tablica 7-9 Liczba reguł o określonej liczbie warunków w poszczególnych kategoriach przed optymalizacją modelu.

| Kategoria | Liczba warunków w regule | | | | | | | | | | | Łączna Liczba reguł |
|----------------------------|--------------------------|----------|-----------|-----------|-----------|-----------|----------|----------|----------|----------|----------|---------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | |
| Znamię łagodne | 0 | 2 | 13 | 16 | 9 | 4 | 3 | 0 | 0 | 0 | 0 | 47 |
| Znamię błękitne | 0 | 0 | 8 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 10 |
| Znamię podejrzane | 1 | 0 | 5 | 11 | 7 | 4 | 3 | 2 | 1 | 4 | 2 | 40 |
| Znamię złośliwe | 0 | 2 | 6 | 3 | 6 | 5 | 1 | 2 | 0 | 0 | 0 | 25 |
| Łączna liczba reguł | 1 | 4 | 32 | 31 | 23 | 13 | 7 | 4 | 1 | 4 | 2 | 122 |

Tablica 7-10 Liczba reguł o określonej liczbie warunków w poszczególnych kategoriach po optymalizację modelu.

| Kategoria | Liczba warunków w regule | | | | | | | | | | | Łączna Liczba reguł |
|----------------------------|--------------------------|----------|-----------|-----------|-----------|-----------|----------|----------|----------|----------|----------|---------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | |
| Znamię łagodne | 0 | 2 | 9 | 18 | 9 | 4 | 1 | 0 | 0 | 0 | 0 | 43 |
| Znamię błękitne | 0 | 0 | 7 | 1 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 10 |
| Znamię podejrzane | 1 | 0 | 6 | 8 | 7 | 7 | 4 | 3 | 0 | 1 | 0 | 37 |
| Znamię złośliwe | 0 | 2 | 7 | 5 | 6 | 1 | 2 | 0 | 0 | 0 | 0 | 23 |
| Łączna liczba reguł | 1 | 4 | 29 | 32 | 24 | 12 | 7 | 3 | 0 | 1 | 0 | 113 |

Zbiór reguł składniowych w kolejnych etapach działania opracowanych algorytmów, zaimplementowanych w opracowanym systemie informatycznym *RuleSEEKER* [Paja, 2005], podlegał – jak stwierdzono – istotnym

modyfikacjom jakościowym, co przedstawiono w Tabelicy 7-11. W tabelicy tej, czerwony kolor czcionki wskazuje reguły usunięte przez operacje generacyjne, natomiast kolorem niebieskim oznaczono reguły dodane do pierwotnego zbioru reguł. Z kolei kolorem zielonym oznaczono reguły, w których zostały zmienione niektóre warunki oraz reguły, powstałe z innych, bardziej szczegółowych. Reguły, które nie uległy żadnej zmianie zapisano czarną czcionką. Numeracja reguł nie jest porównywalna, co jest spowodowane procesami sortowania reguł w środowisku systemu *RuleSEEKER*.

Tablica 7-11. Przykład modyfikacji zbioru reguł decyzji uzyskanego poprzez analizę bazy informacyjnej P548134 z użyciem algorytmu GTS. W bazie zastosowano skróty nazw atrybutów i ich wartości w celu przedstawienia wyników w czytelnej formie. Pełny opis zawartości i charakterystykę bazy zamieszczono w Załączniku A.

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|---|---|
| <p>REGUŁA 1 JEŻELI D_CIAŁbarw JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 2 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 3 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BŁĘKITNY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 13 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BŁĘKITNY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 4 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> | <p>REGUŁA 1 JEŻELI D_CIAŁbarw JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 2 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 3 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 2.5 ORAZ BRZEG < 4.5 ORAZ C_BŁĘKITNY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 4 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|--|
| <p>REGUŁA 5 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 6 JEŻELI C_BŁĘKITNY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 7 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ D_CIAŁbarw JEST nie_ma ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 8 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_CZARNY JEST jest ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 9 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_CZARNY JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 10 JEŻELI BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_CZERWONY JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 11 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 12 JEŻELI C_BŁĘKITNY JEST nie_ma ORAZ D_POLEbezs JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 14</p> | <p>REGUŁA 5 JEŻELI C_BŁĘKITNY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 6 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ D_CIAŁbarw JEST nie_ma ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 7 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_CZARNY JEST jest ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 8 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_CZARNY JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 9 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 10 JEŻELI C_BŁĘKITNY JEST nie_ma ORAZ D_POLEbezs JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 11</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|--|
| <p>JEŻELI C_BIAŁY JEST jest ORAZ C_J_BRAZOWY JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 15 JEŻELI C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 16 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 17 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 18 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 19 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 20 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 21 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ D_KROPbarw JEST nie_ma</p> | <p>JEŻELI C_BIAŁY JEST jest ORAZ C_J_BRAZOWY JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 12 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 13 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 14 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 15 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BŁĘKITNY JEST nie_ma</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|---|
| <p>ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 22 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 23 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 24 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 25 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 26 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAZOWY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 27 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 28 JEŻELI BRZEG >= 0</p> | <p>ORAZ D_KROPbarw JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 16 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 17 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 18 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAZOWY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 19 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 20 JEŻELI BRZEG >= 0</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|--|
| <p>ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 29 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST nie_ma ORAZ C_CZARNY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 30 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 31 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZERWONY JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 32 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 33 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 34 JEŻELI BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma</p> | <p>ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 21 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST nie_ma ORAZ C_CZARNY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 22 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 23 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZERWONY JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 24 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 25 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|---|
| <p>ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 35 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 36 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_BŁĘKITNY JEST jest ORAZ C_C_BRAZOWY JEST jest ORAZ C_J_BRAZOWY JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 37 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 38 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 39 JEŻELI C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST nie_ma ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZERWONY JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 40 JEŻELI C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZERWONY JEST jest</p> | <p>REGUŁA 26 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 27 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_BŁĘKITNY JEST jest ORAZ C_J_BRAZOWY JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 28 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 29 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 30 JEŻELI C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST nie_ma ORAZ C_J_BRAZOWY JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 31 JEŻELI C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZERWONY JEST jest</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|---|
| <p>ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 41 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 42 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 43 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 44 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 45 JEŻELI BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZARNY JEST nie_ma ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> | <p>ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 32 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 33 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|---|
| <p>REGUŁA 46 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 47 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> | <p>REGUŁA 34 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_CZERWONY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 35 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 36 JEŻELI C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST nie_ma ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 37 JEŻELI BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAZOWY JEST nie_ma ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 38 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAZOWY JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 39 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> |

| <p>Wtórne źródło wiedzy (przed optymalizacją) (1)</p> | <p>Wtórne źródło wiedzy (po optymalizacji) (2)</p> |
|--|---|
| | <p>REGUŁA 40 JEŻELI BRZEG >= 7.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_C_BRAZOWY JEST jest ORAZ C_J_BRAZOWY JEST jest ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 41 JEŻELI C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 42 JEŻELI BRZEG >= 7.5 ORAZ C_C_BRAZOWY JEST jest ORAZ C_CZARNY JEST nie_ma ORAZ D_CIAŁbarw JEST jest ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> <p>REGUŁA 43 JEŻELI BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_łagodna</p> |
| <p>REGUŁA 48 JEŻELI BRZEG >= 0.5 ORAZ BRZEG < 1.5 TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 49 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BIAŁY JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 50 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_BIAŁY JEST jest ORAZ D_ROZGpasm JEST nie_ma</p> | <p>REGUŁA 44 JEŻELI BRZEG >= 0.5 ORAZ BRZEG < 1.5 TO DIAGNOZA JEST Mel_podejrz</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją) (1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji) (2)</p> |
|--|---|
| <p>TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 51 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_CZARNY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 52 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ D_CIAŁbarw JEST nie_ma ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 53 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_J_BRAZOWY JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 54 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_CIAŁbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 55 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_BIAŁY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 56 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ D_CIAŁbarw JEST nie_ma ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 57 JEŻELI ASYMETRIA JEST As_2_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_CZARNY JEST nie_ma</p> | <p>REGUŁA 45 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_J_BRAZOWY JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 46 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_CIAŁbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 47 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 48 JEŻELI ASYMETRIA JEST As_2_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_CZARNY JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|---|--|
| <p>TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 58 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_BIAŁY JEST jest ORAZ C_CZARNY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 59 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_CZERWONY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 60 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 61 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_CZERWONY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 62 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 63 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_BŁĘKITNY JEST jest ORAZ C_C_BRAŻOWY JEST nie_ma ORAZ C_J_BRAŻOWY JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 64 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5</p> | <p>REGUŁA 49 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 50 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_CZERWONY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 51 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 52 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_BŁĘKITNY JEST jest ORAZ C_C_BRAŻOWY JEST nie_ma ORAZ C_J_BRAŻOWY JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|---|
| <p>ORAZ BRZEG < 8 ORAZ D_KROPbarw JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 65 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 66 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST jest ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZARNY JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 67 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_POLEbezs JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 68 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ C_CZARNY JEST nie_ma ORAZ C_CZERWONY JEST jest ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 69 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_J_BRAZOWY JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 70 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 6.5 ORAZ BRZEG < 7.5</p> | <p>REGUŁA 53 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 54 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST jest ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZARNY JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 55 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_POLEbezs JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 56 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_J_BRAZOWY JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 57 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_CZARNY JEST nie_ma</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|--|
| <p>ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 71 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ C_BŁĘKITNY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST nie_ma ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 72 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 73 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_C_BRAŻOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 74 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BIAŁY JEST jest ORAZ C_C_BRAŻOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 75 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BIAŁY JEST jest ORAZ C_J_BRAŻOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 76</p> | <p>ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 58 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ C_BŁĘKITNY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST nie_ma ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 59 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 60 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_C_BRAŻOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 61 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BIAŁY JEST jest ORAZ C_J_BRAŻOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 62</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|---|---|
| <p>JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 77 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST jest ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_CIAŁbarw JEST jest ORAZ D_POLEbezs JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 78 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST jest ORAZ C_C_BRAZOWY JEST jest ORAZ C_CZARNY JEST jest ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 79 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST jest ORAZ C_CZARNY JEST nie_ma ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 80 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST jest ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAZOWY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_POLEbezs JEST jest</p> | <p>JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_CZERWONY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 63 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST jest ORAZ C_CZARNY JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 64 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 65 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST jest ORAZ C_C_BRAZOWY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|---|
| <p>ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 81 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRĄZOWY JEST jest ORAZ C_J_BRĄZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 82 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRĄZOWY JEST jest ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 83 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_C_BRĄZOWY JEST jest ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 84 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRĄZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest</p> | <p>TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 66 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_C_BRĄZOWY JEST jest ORAZ C_J_BRĄZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 67 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 68 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ C_CZERWONY JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|---|--|
| <p>ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 85 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 86 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAŻOWY JEST jest ORAZ C_J_BRAŻOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 87 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_C_BRAŻOWY JEST jest ORAZ C_J_BRAŻOWY JEST jest ORAZ C_CZARNY JEST nie_ma ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> | <p>REGUŁA 69 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 7.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_C_BRAŻOWY JEST nie_ma ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 70 JEŻELI BRZEG >= 7.5 ORAZ C_BŁĘKITNY JEST jest ORAZ C_C_BRAŻOWY JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 71 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ C_BIAŁY JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 72 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_J_BRAŻOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> |

| <p>Wtórne źródło wiedzy (przed optymalizacją)</p> <p>(1)</p> | <p>Wtórne źródło wiedzy (po optymalizacji)</p> <p>(2)</p> |
|--|--|
| | <p>REGUŁA 73 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST nie_ma ORAZ C_CZARNY JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 74 JEŻELI BRZEG >= 7.5 ORAZ D_KROPbarw JEST nie_ma ORAZ D_POLEbezs JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 75 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 76 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 7.5 ORAZ C_BIAŁY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 77 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 5.5 ORAZ C_BIAŁY JEST jest ORAZ C_C_BRAŻOWY JEST jest ORAZ C_J_BRAŻOWY JEST jest ORAZ D_KROPbarw JEST nie_ma ORAZ D_POLEbezs JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 78 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 7.5 ORAZ C_J_BRAŻOWY JEST jest ORAZ C_CZERWONY JEST jest</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|--|
| | <p>ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 79 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ C_BIAŁY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_podejrz</p> <p>REGUŁA 80 JEŻELI BRZEG >= 7.5 ORAZ C_CZARNY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_podejrz</p> |
| <p>REGUŁA 88 JEŻELI ASYMETRIA JEST As_2_osiowa ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 89 JEŻELI ASYMETRIA JEST As_2_osiowa ORAZ C_CZARNY JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 90 JEŻELI C_BIAŁY JEST jest ORAZ C_BŁĘKITNY JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 91 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_BIAŁY JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 92 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_BIAŁY JEST jest ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> | <p>REGUŁA 81 JEŻELI ASYMETRIA JEST As_2_osiowa ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 82 JEŻELI ASYMETRIA JEST As_2_osiowa ORAZ C_CZARNY JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 83 JEŻELI C_BIAŁY JEST jest ORAZ C_BŁĘKITNY JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 84 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_BIAŁY JEST jest ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 85 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_BIAŁY JEST jest ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|---|--|
| <p>REGUŁA 93 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ D_CIAŁbarw JEST nie_ma TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 94 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_C_BRAZOWY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 95 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_CZARNY JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 96 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_BIAŁY JEST jest ORAZ C_C_BRAZOWY JEST jest ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 97 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_BIAŁY JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 98 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_BIAŁY JEST nie_ma ORAZ D_POLEbezs JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 99 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_CZERWONY JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST jest</p> | <p>REGUŁA 86 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ D_CIAŁbarw JEST nie_ma TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 87 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_C_BRAZOWY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 88 JEŻELI BRZEG >= 6.5 ORAZ BRZEG < 7.5 ORAZ C_CZARNY JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 89 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_BIAŁY JEST jest ORAZ C_C_BRAZOWY JEST jest ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 90 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_BIAŁY JEST nie_ma ORAZ D_POLEbezs JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 91 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_CZERWONY JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST jest</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|---|---|
| <p>TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 100 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 101 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_BIAŁY JEST jest ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZARNY JEST jest ORAZ D_POLEbezs JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 102 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 103 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BIAŁY JEST jest ORAZ C_J_BRAZOWY JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 104 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 105 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BIAŁY JEST nie_ma ORAZ D_CIAŁbarw JEST jest</p> | <p>TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 92 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_BIAŁY JEST jest ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZARNY JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 93 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_J_BRAZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 94 JEŻELI BRZEG >= 2.5 ORAZ BRZEG < 3.5 ORAZ C_BIAŁY JEST jest ORAZ C_J_BRAZOWY JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 95 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ C_BIAŁY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|---|---|
| <p>ORAZ D_POLEbezs JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 106 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_C_BRAZOWY JEST jest ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 107 JEŻELI BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ C_CZARNY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 108 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_CZERWONY JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 109 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 110 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 2.5</p> | <p>REGUŁA 96 JEŻELI BRZEG >= 5.5 ORAZ BRZEG < 6.5 ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 97 JEŻELI BRZEG >= 3.5 ORAZ BRZEG < 4.5 ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 98 JEŻELI C_BŁĘKITNY JEST jest ORAZ C_CZARNY JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 99 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 2.5</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|--|--|
| <p>ORAZ BRZEG < 3.5 ORAZ C_CZARNY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 111 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ BRZEG < 8 ORAZ C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST nie_ma ORAZ C_CZARNY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 112 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ C_CZARNY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> | <p>ORAZ BRZEG < 3.5 ORAZ C_CZARNY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 100 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ D_CIAŁbarw JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 101 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ C_CZARNY JEST jest ORAZ D_KROPbarw JEST jest ORAZ D_ROZGpasm JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 102 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ BRZEG >= 7.5 ORAZ D_CIAŁbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> <p>REGUŁA 103 JEŻELI ASYMETRIA JEST As_1_osiowa ORAZ C_J BRAZOWY JEST jest ORAZ C_CZERWONY JEST jest ORAZ D_CIAŁbarw JEST jest ORAZ D_POLEbezs JEST jest ORAZ D_ROZGpasm JEST jest ORAZ D_SIEĆbarw JEST jest TO DIAGNOZA JEST Mel_złośliw</p> |
| <p>REGUŁA 113 JEŻELI C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST jest ORAZ C_CZARNY JEST nie_ma</p> | <p>REGUŁA 104 JEŻELI C_BIAŁY JEST nie_ma ORAZ C_BŁĘKITNY JEST jest ORAZ C_CZARNY JEST nie_ma</p> |

| <p style="text-align: center;">Wtórne źródło wiedzy (przed optymalizacją)</p> <p style="text-align: center;">(1)</p> | <p style="text-align: center;">Wtórne źródło wiedzy (po optymalizacji)</p> <p style="text-align: center;">(2)</p> |
|---|--|
| <p>TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 114 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST jest ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 115 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_BŁĘKITNY JEST jest ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 116 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 117 JEŻELI C_BŁĘKITNY JEST jest ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZARNY JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 118 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_CZARNY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 119 JEŻELI C_BŁĘKITNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 120 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ BRZEG >= 4.5 ORAZ BRZEG < 5.5 ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 121 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma</p> | <p>TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 105 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_BŁĘKITNY JEST jest ORAZ D_KROPbarw JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 106 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_BŁĘKITNY JEST jest ORAZ D_KROPbarw JEST jest TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 107 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_BŁĘKITNY JEST jest ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 108 JEŻELI C_BŁĘKITNY JEST jest ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZARNY JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 109 JEŻELI BRZEG >= 1.5 ORAZ BRZEG < 2.5 ORAZ C_CZARNY JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 110 JEŻELI C_BŁĘKITNY JEST jest ORAZ C_CZERWONY JEST nie_ma ORAZ D_CIAŁbarw JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne</p> <p>REGUŁA 111 JEŻELI BRZEG >= 0 ORAZ BRZEG < 0.5 ORAZ C_J_BRAZOWY JEST nie_ma ORAZ C_CZERWONY JEST nie_ma</p> |

| Wtórne źródło wiedzy (przed optymalizacją) (1) | Wtórne źródło wiedzy (po optymalizacji) (2) |
|--|---|
| ORAZ D_CIAŁbarw JEST jest TO DIAGNOZA JEST Mel_znmbłękitne REGUŁA 122 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_C_BRAZOWY JEST jest ORAZ C_CZARNY JEST jest ORAZ D_KROPbarw JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne | ORAZ D_CIAŁbarw JEST jest TO DIAGNOZA JEST Mel_znmbłękitne REGUŁA 112 JEŻELI ASYMETRIA JEST Zmiana_sym ORAZ C_C_BRAZOWY JEST jest ORAZ C_CZARNY JEST jest ORAZ D_KROPbarw JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne REGUŁA 113 JEŻELI C_BIAŁY JEST nie_ma ORAZ C_CZARNY JEST nie_ma ORAZ D_KROPbarw JEST nie_ma ORAZ D_ROZGpasm JEST nie_ma ORAZ D_SIEĆbarw JEST nie_ma TO DIAGNOZA JEST Mel_znmbłękitne |

Omawiając etap usunięcia reguł nieużywanych, należy podkreślić, że wszystkie badane algorytmy generują nadmiarowe reguły. Operacja usunięcia reguł nieużywanych (*NU*) eliminuje 15 zbędnych reguł (Tablica 7-1 oraz 7-11), które nie pokrywają żadnego przypadku analizowanej bazy informacyjnej. Kolejnym elementem optymalizacji jest operacja usunięcia warunków zbędnych (*W*), co pozwala ograniczyć model o 40 nadmiarowych warunków, bez zmiany jego parametrów. Następną wykonywaną operacją jest łączenie reguł (*L*), w wyniku czego cztery reguły (3, 13 oraz 41, 44; Tablica 7-11, kolumna 1) łączone są w dwie bardziej ogólne reguły (3 oraz 32; Tablica 7-11, kolumna 2). W efekcie powyższych operacji usunięcia nadmiarowości, otrzymano model uczenia zawierający 105 reguł (Tablica 7-1), cechujący się błędem klasyfikacji równym 5.3%. Następną operacją, polegającą na dodaniu nowych reguł (31 reguł) w celu utworzenia modelu uczenia cechującego się możliwie niewielkim błędem klasyfikacji, została uznana za priorytetowy cel optymalizacji zanalizowanego modelu: błąd klasyfikacji osiągnął nawet 0%. Końcową operacją (według zakładanej kolejności) jest wybór reguł finalnych (*FR*), pozwalający usunąć 17 reguł.

Niekiedy celowym może być wykonanie drugiej serii wspomnianych operacji, w celu ponownej optymalizacji utworzonego uprzednio zbioru reguł. W rozpatrywanym przykładzie zastosowano dodatkowo operacje usunięcia warunków zbędnych (118 warunków), połączenia reguł (2 reguły) oraz ponowny wybór reguł finalnych. W efekcie utworzony został model składający się ze 113 reguł, cechujący się błędem klasyfikacji równym 0%. Ponadto, średnia dokładność reguł wynosi 1, co oznacza, że żadna reguła nie klasyfikowała błędnie przypadków z analizowanej bazy informacyjnej.

Analogiczną analizę prowadzono w odniesieniu do wszystkich badanych baz informacyjnych. Zamieszczenie i omówienie wszystkich modeli byłoby wręcz niemożliwe (tutaj jeden model opisano na 22 stronach, jednak przesłedzenie go pokazuje ogólną metodykę analizy modeli).

Podsumowując, analiza uzyskanych wyników klasyfikacji wskazuje, że zastosowanie algorytmu optymalizacji wtórnych źródeł wiedzy (modeli uczenia) umożliwia uzyskanie znacznego zmniejszenia wartości błędu klasyfikacji, bez względu na typ analizowanych danych. Na przykład, analizując bazę znamion melanocytowych skóry, przedział błędu klasyfikacji przed optymalizacją (2.4%-6.9%) zmniejszył się do 0%-3.1%. Analogiczny skutek, tj. zmniejszenie przedziału błędu klasyfikacji został osiągnięty podczas analizy bazy informacyjnej *GLASS*. Wyniki optymalizacji modeli globalnych dla obydwu opisywanych tu baz informacyjnych, spowodowały zmniejszenie błędu klasyfikacji do 0.7% (baza *P548134*) oraz 2.8% (baza *GLASS*). Wydaje się zatem, że opracowana metodologia budowy optymalnych modeli uczenia przy użyciu różnych algorytmów uczenia maszynowego na podstawie tej samej bazy informacyjnej przynosi zamierzony skutek wyrażony w postaci polepszenia parametrów charakteryzujących generowane modele uczenia. Ponadto wyniki badań sugerują, że można osiągnąć bardziej korzystne wyniki klasyfikacji wskutek optymalizacji modelu globalnego, tzn. modelu stanowiącego stop modeli składowych, uzyskanych

przy pomocy różnych algorytmów uczenia maszynowego. Wydaje się, że uzyskane wyniki dodatkowo, jakby w sposób niezamierzony, potwierdziły słuszność paradygmatu sztucznej inteligencji (Rozdział 1, strona 5), stanowiącego punkt wyjścia niniejszej rozprawy.

ROZDZIAŁ 8

PODSUMOWANIE

Badania zrealizowane w ramach niniejszej rozprawy zostały zakończone opracowaniem i implementacją nowej metodologii optymalizacji wielu modeli uczenia, uzyskanych na podstawie tej samej tablicy decyzji.

Podstawowy nurt przeprowadzonych badań dotyczył zdefiniowania algorytmu optymalizującego modele uczenia w postaci zbioru reguł decyzji. Do tego celu zastosowano istniejące i znane algorytmy uczenia maszynowego: *GTS*, *LEM2* oraz *ID3*. W toku badań zdefiniowano siedem odmiennych operacji możliwych do wykonania na zbiorze reguł i do implementacji programowej. Opracowany algorytm wymagał zaprojektowania i utworzenia profesjonalnego narzędzia informatycznego (system informatyczny *RuleSEEKER*). Modułowa budowa wspomnianego narzędzia umożliwia łatwą modyfikację zaimplementowanych operacji oraz dodanie nowych rozwiązań.

Ostatnim elementem pracy były klasyfikacyjne badania porównawcze opracowanych algorytmów. Badania zostały przeprowadzone w procesie analizy bazy informacyjnej znamion melanocytowych skóry, a także w odniesieniu do baz informacyjnych zaczerpniętych z repozytorium Uniwersytetu Kalifornijskiego w Irvine, opisanych w Załączniku A. Badane bazy informacyjne były zróżnicowane pod względem rodzaju atrybutów i ich liczebności. Analiza uzyskanych wyników klasyfikacji, przedstawiona

w Rozdziale 7, wskazuje, że zastosowanie algorytmu optymalizacji wtórnych źródeł wiedzy umożliwia uzyskanie znacznego zmniejszenia wartości błędu klasyfikacji jak i poprawy parametrów opisujących badane modele uczenia, bez względu na typ analizowanych danych. Ponadto, modele podlegają istotnym modyfikacjom jakościowym. Następuje pewne uogólnienie modelu uczenia, reguły stają się bardziej ogólne, przy jednoczesnym zmniejszeniu ich liczby. Wykazano również, iż połączenie modeli w jeden model globalny a następnie jego optymalizacja pozwala uzyskać modele o właściwościach porównywalnych a niekiedy lepszych w stosunku do odrębnie optymalizowanych modeli. Zastosowane operacje generyczne znacznie eliminują zjawisko nadmiarowości w modelach uczenia, uzupełniając je nowymi, brakującymi regułami.

Przeprowadzone badania potwierdzają spełnienie tezy, że różne systemy uczenia maszynowego, zastosowane do analizy tych samych danych, mogą tworzyć modele uczenia o odmiennych właściwościach. Ponadto możliwe jest opracowanie metodologii generowania na ich podstawie modeli uczenia, cechujących się analogiczną, lub lepszą zdolnością klasyfikacji oraz identyfikacji nieznanymi obiektów, opisanych w logice atrybutalnej

ZAŁĄCZNIK **A**

OPIS BADANYCH BAZ INFORMACYJNYCH

Opracowane algorytmy optymalizacji wtórnych źródeł wiedzy zostały prze-testowane stosując dane, pochodzące z repozytorium baz danych dla uczenia maszynowego Uniwersytetu Kalifornijskiego w Irvine [Newman i inni, 1998]. Głównym jednak obiektem badań była własna baza informacyjna o zagrożeniach czerniakiem skóry ludności południowo-wschodniej Polski [Hippe i inni, 2003], zwaną dalej bazą znamion melanocytowych skóry. Ilościowa informacja na temat tych baz została zamieszczona w Tablicy A-1. Dodatkowo, w dalszej części rozdziału omówiono najważniejszą dla weryfikacji głównego celu moich badań, bazę znamion melanocytowych skóry.

Tablica A-1 Charakterystyka baz informacyjnych użytych do testowania opracowanych algorytmów optymalizacji wtórnych źródeł wiedzy (kursywą podano skrócone nazwy baz, które używane są w tekście).

| Oryginalna nazwa bazy <i>Skrócona nazwa bazy</i> | Treść bazy | Liczba atrybutów | Liczba przypadków | Liczba klas |
|--|--|-------------------------|--------------------------|--------------------|
| GLASS IDENTIFICATION <i>GLASS</i> | Opis składu chemicznego różnych rodzajów szkła [German, 1987]. | 9 | 214 | 6 |

| Oryginalna nazwa bazy <i>Skrócona nazwa bazy</i> | Treść bazy | Liczba atrybutów | Liczba przypadków | Liczba klas |
|--|---|------------------|----------------------|-------------|
| WINE RECOGNITION <i>WINE</i> | Dane zawierające informację na temat gatunków włoskich win na podstawie analizy chemicznej [Forina, 1991]. | 13 | 178 | 3 |
| CAR EVALUATION DATABASE <i>CAR</i> | Klasyfikacja samochodów według parametrów [Bohanec i Rajkovic, 1988] | 6 | 998 (ograniczona) | 4 |
| HEART-DISEASE <i>HEART</i> | Przypadki diagnozy chorób serca [Detrano i inni, 1989] | 14 | 270 | 5 |
| BAZA ZNAMION MELANOCYTO- WYCH SKÓRY <i>P548134</i> | Opis znamion malanocytowych skóry; bardziej szczegółową charakterystykę bazy podano w dalszej treści rozdziału. | 13 | 548 | 4 |

Baza zmian malanocytowych skóry

Baza informacyjna *P548134* zawiera statystycznie znaczącą liczbę (548) zweryfikowanych histologicznie przypadków, należących do jednego z czterech¹⁾ rodzajów znamion melanocytowych skóry: (i) <znamię łagodne>, (ii) <znamię błękitne>, (iii) <znamię podejrzone> oraz (iv) <znamię złośliwe>. Przypadki opisano przy pomocy trzynastu atrybutów. Atrybuty

¹⁾ Czasami w literaturze wymienia się jedynie trzy kategorie znamion melanocytowych skóry, pomijając <znamię błękitne>. Znamię to jest wówczas traktowane jako pod-kategoria znamienia łagodnego.

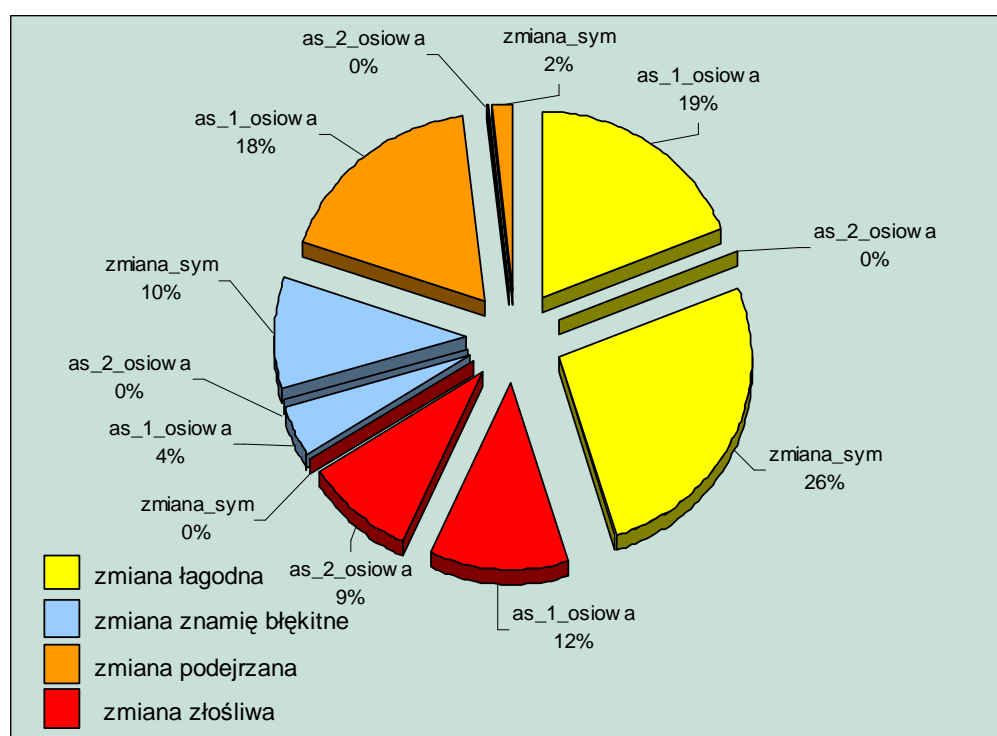
opisujące przypadki ujęte w omawianej bazie informacyjnej mogą być formalnie przypisane do czterech grup zmiennych, tworzących znaną w dermatologii regułę **ABCD** {**A**symmetry, **B**order, **C**olor, **D**iversity of structures} [Stolz i inni, 1993]. W regule tej (w odniesieniu do polskiej wersji językowej) **A** <Asymetria> – oznacza asymetrię znamienia, **B** <Brzeg> – informuje o charakterze obrzeża znamienia, **C** <Kolor> – zawiera informację o liczbie i rodzaju barw znamienia, natomiast – **D** <Struktura> – mówi o liczbie i charakterze zróżnicowania dopuszczalnych struktur znamienia. Wartości logiczne poszczególnych atrybutów opisujących dowolny przypadek w omawianej bazie przedstawia Tablica A-2.

Tablica A-2 Wartości logiczne atrybutów opisujących znamiona melanocytowe skóry

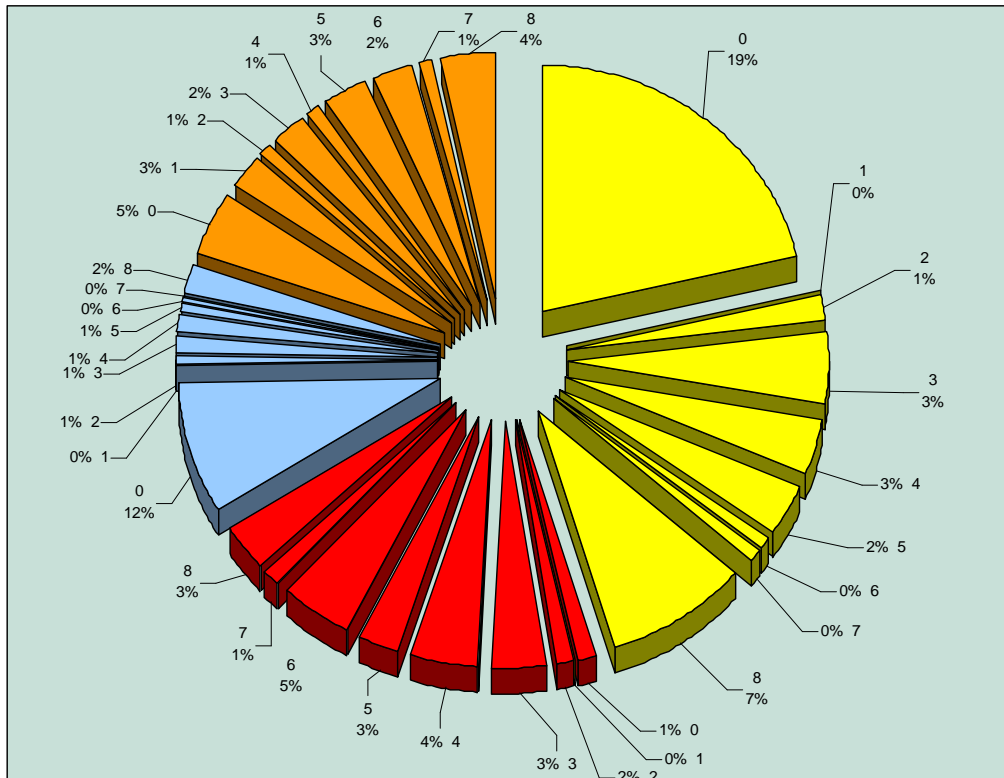
| Atrybut | Wartość logiczna |
|----------------------------|--|
| Asymetria | zmiana symetryczna asymetria 1-osiowa asymetria 2-osiowa |
| Brzeg | wartość liczbowa z przedziału 0 ÷ 8 |
| Kolor | |
| - biały | jest / nie ma |
| - błękitny | jest / nie ma |
| - ciemnobrązowy | jest / nie ma |
| - jasnobrązowy | jest / nie ma |
| - czarny | jest / nie ma |
| - czerwony | jest / nie ma |
| Struktura | |
| - ciałka barwnikowe | jest / nie ma |
| - kropki barwnikowe | jest / nie ma |
| - pole bezstrukturalne | jest / nie ma |
| - rozgałęzienia pasmowate | jest / nie ma |
| - sieć barwnikowa | jest / nie ma |
| Kategoria znamienia | znamie łagodne znamie błękitne |

| | |
|--|-------------------|
| | znamie podejrzane |
| | znamie zlosliwe |

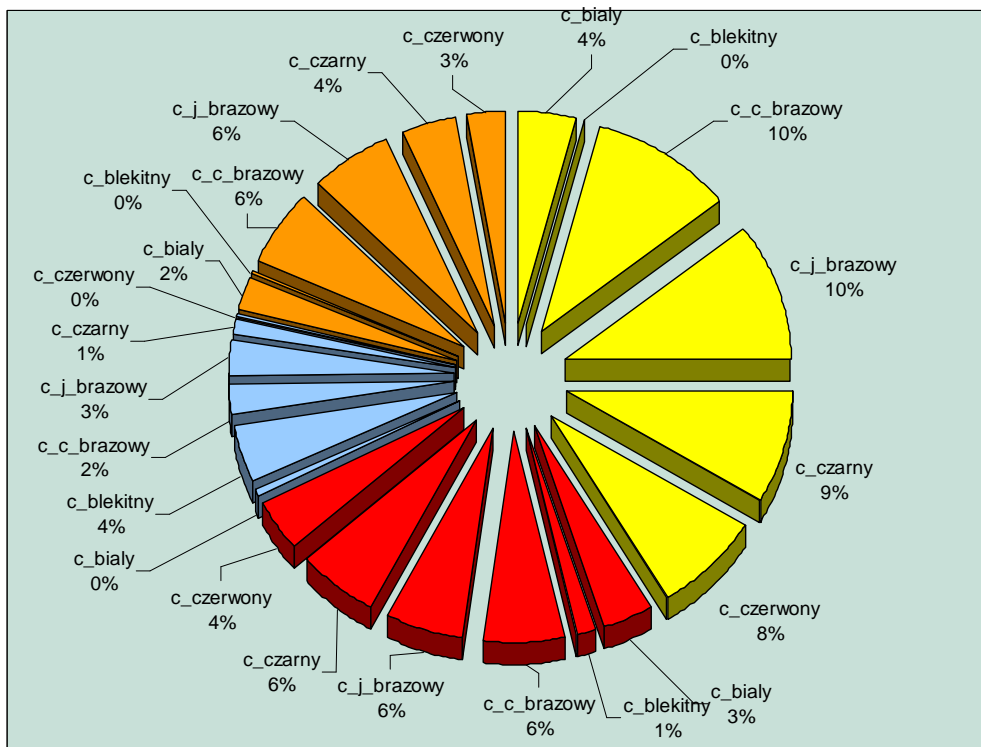
Badana baza informacyjna jest pozbawiona błędów, dzięki weryfikacji jej treści przy pomocy specjalnie opracowanych szablonów. Analiza statystyczna treści bazy danych dała wyniki przedstawione na Rys. A-1 do A-4. Na wykresach kołowych przedstawiono procentowy udział wartości atrybutu rozpatrywanego w znamieniu danego typu.



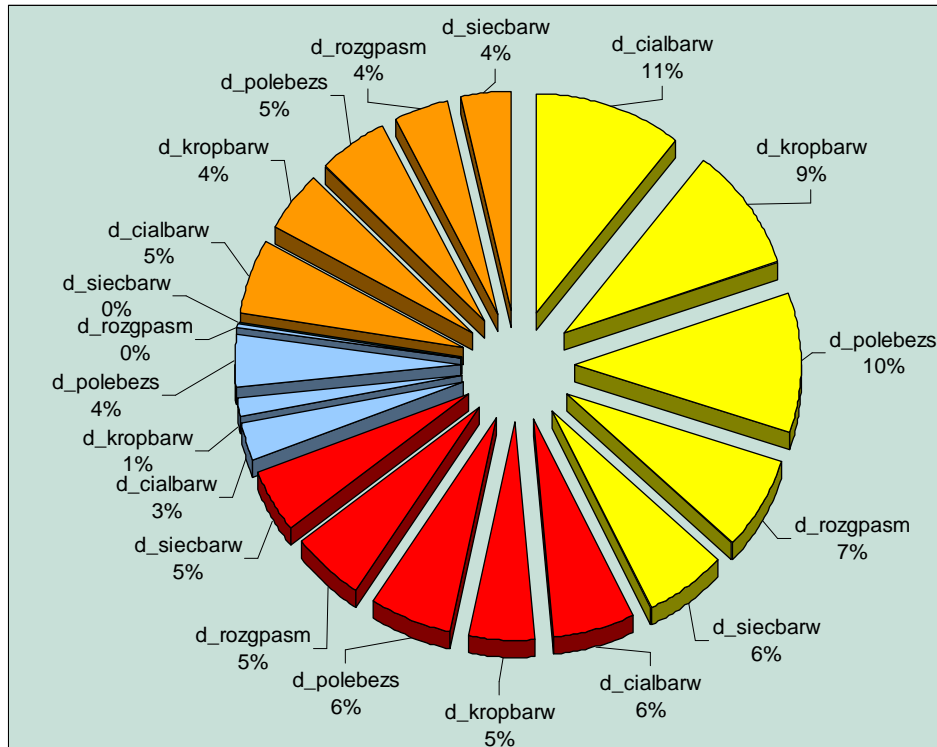
Rys. A-1 Udział atrybutu <Asymetria> w rozważanych kategoriach znamion melanocytowych



Rys. A-2 Udział atrybutu <Brzeg> w rozważanych kategoriach znamion melanocytowych



Rys. A-3 Udział atrybutu <Kolor> w rozważanych kategoriach znamion melanocytowych



Rys. A-4 Udział atrybutu <Struktura> w rozważanych kategoriach znamion melano-
cytowych

LITERATURA

[AI-TECH, 2000]

AI-TECH: www.aitech.com.pl

[Ali i Pazzani, 1993] K. Ali, M. Pazzani:

HYDRA: A noise-tolerant relational concept learning algorithm

W: Proc. 13th Int. Joint Conference on Artificial Intelligence (IJCAI'93), Morgan Kaufmann, Chambéry (France) 1993.

[An, 2003] A. An:

Learning Classification Rules from Data

W: Int. Journal of Computers and Mathematics with Applications, **45**(2003, No 4-5)737-748.

[An i Wang, 2001] A. An, Y. Wang:

Comparisons of Classification Methods for Screening Potential Compounds

W: Proceedings IEEE International Conference on Data Mining, ICDM 2001, San Jose, CA, USA, s. 11-18.

[An i Cercone, 2001] A. An, N. Cercone:

Rule quality measures for rule induction systems: description and evaluation

W: Computational Intelligence, Blackwell Publishers, **17**(2001, No 3), Malden (USA), Oxford (UK) 2001.

[Bazan, 1998] J. Bazan:

A comparison of dynamic and non-dynamic rough set methods for extracting laws from decision tables

W: Polkowski L., Skowron A. (red.), Rough sets in knowledge discovery, vol 1, Methodology and applications, Physica-Verlag, Heidelberg 1998, s. 321-365.

[Bishop, Fienberg oraz Holand, 1991] Y.M.G. Bishop, S.E. Fienberg, P.W. Holand:

Discrete Multivariate Analysis: Theory and Practice

The MIT Press, 1991.

[Bohanec i Rajkovic, 1988] M. Bohanec, V. Rajkovic:

Knowledge acquisition and explanation for multi-attribute decision making

W: 8th Int. Workshop on Expert Systems and their Applications, Avignon, Franc 1988, s. 59-78.

[Bonner, Galant i Tyburcy, 2001] R.F. Bonner, V. Galant, J. Tyburcy:

A Class of Lazy Learning Algorithms and Knowledge Acquisition from Databases

W: Boborski A.J., Bonner R.F., Owoc M.L. (Eds.), Knowledge Acquisition and Distributed Learning in Resolving Managerial Issues, Mälardalen University Press, Västerås (Sweden) 2001.

[Bramer, 2000] M.A. Bramer:

Automatic Induction of Classification Rules from Examples Using N-Prism

W: Research and Development in Intelligent Systems XVI. Springer-Verlag, 2000, s. 99-121.

[Brazdil oraz Torgo, 1990] P. Brazdil, L. Torgo:

Knowledge Acquisition via Knowledge Integration

In: Current Trends in Knowledge Acquisition, IOS Press, 1990.

[Breiman i inni, 1984] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone:

Classification and Regression Trees

Belmont Wadsworth Int. Group, New York 1984.

[Brijs, Vanhoof i Wets, 2000] T. Brijs, K. Vanhoof, G. Wets:

Reducing redundancy in characteristic rule discovery by using integer programming techniques

W: Intelligent Data Analysis Journal, IOS Press, 4(2000)229-240.

[Bruha oraz Kockova, 1993] I. Bruha, S. Kockova:

Quality of Decision Rules: Empirical and Statistical Approaches

Informatica, 17(1993)233-243.

[Carvalho i Freitas, 2004] D.R. Carvalho, A.A. Freitas:

New results for a hybrid decision tree/genetic algorithm for data mining

W: J. Garibaldi, A. Lofti, R. John, (Eds.), Proc. 4th Int. Conference on Recent Advances in Soft Computing (RASC-2002), Nottingham Trent University 2002, s. 260-265.

[Cendrowska, 1987] J. Cendrowska:

PRISM: An algorithm for inducing modular rules

W: Int. J. Man-Machine Studies 27(1987) 349-370.

[Cestnik, 1990] B. Cestnik:

Estimating probabilities: a crucial task in machine learning

W: Proc. ECAO 90, Stockholm (Sweden) 1990.

[Cichosz, 2000] P. Cichosz:

Systemy uczące się

Wydawnictwa Naukowo-Techniczne, Warszawa 2000, s. 41-44.

[Clark i Boswell, 1991] P. Clark, R. Boswell:

Rule induction with CN2: some recent improvements

W: Kodratoff Y. (Ed.), Proc. 5th European Working Session on Learning (EWSL'91), Springer-Verlag, Porto (Portugal) 1991, s. 151-163.

- [Clark i Niblett, 1989] P. Clark, T. Niblett:
The CN2 Induction Algorithm
 Machine Learning Journal, Kluwer Academic Publishers, **3**(1989)261-283, Netherlands 1989.
- [Cohen, 1960] J. Cohen:
A Coefficient of Agreement for Nominal Scales
 Educational and Psych Meas., **22**(1960)37-46.
- [Cragun i Steudel, 1987] B.J. Cragun, H.J. Steudel:
A Decision-Table-Based Processor for Checking Completeness and Consistency in Rule-Based Expert Systems
 W: Int. Journal of Man-Machine Studies, 26(5)1987, s. 633-648.
- [Detrano i inni, 1989] R. Detrano, A. Janosi, W. Steinbrunn, M. Pfisterer, J. Schmid, S. Sandhu, K. Guppy, S. Lee, V. Froelicher:
International application of a new probability algorithm for the diagnosis of coronary artery disease
 American Journal of Cardiology, **64**(1989)304-310.
- [Domingos, 1995] P. Domingos:
Two-way induction
 Proc. 7th IEEE International Conference on Tools with Artificial Intelligence, IEEE Computer Society Press, Herndon (USA) 1995, s. 182-189.
- [Domingos, 1996] P. Domingos:
Unifying Instance-Based and Rule-Based Induction
 Machine Learning, Kluwer Academic Publishers, **24**(1996)141-168, Boston (USA) 1996.
- [Duch, Adamczak i Grąbczewski, 1996] W. Duch, R. Adamczak, K. Grąbczewski:
Extraction of logical rules from training data using backpropagation networks
 CAI'96, 1st Polish Conference on Theory and Applications of Artificial Intelligence, 19-21.12.1996, Łódź, s. 171-178.
- [Duch, Adamczak i Grąbczewski, 1999] W. Duch, R. Adamczak, K. Grąbczewski:
Methodology of extraction, optimization and application of logical rules
 Intelligent Information Systems VIII, Ustroń (Poland) 1999, s. 22-31.
- [Duda, Hart i Stork, 2001] R.O. Duda, P.E. Hart, D.G. Stork:
Pattern Classification
 John Wiley & Sons, Inc. Chichester 2001, s.12-17, 431-435.
- [Duda, Gaschnig i Hart, 1979] R. Duda, J. Gaschnig, P. Hart:
Model design in the prospector consultant system for mineral exploration
 W: Michie D., (Ed.) Expert Systems in the Micro-electronic Age, University Press, Edinburgh (UK) 1979.

[Eijkel, 1999] G. Eijkel:

Rule induction

W: Berthold M., Hand D. (Eds.), *Intelligent Data Analysis. An Introduction*, Springer-Verlag, Berlin 1999, s. 195-216.

[Feng i Michie, 1994] C. Feng, D. Michie:

Machine Learning of Rules and Trees

W: D. Michie, D.J. Spiegelhalter, C.C. Taylor, (Eds.) *Machine Learning, Neural and Statistical Classification*, Ellis Horwood, London, New York 1994, s. 77-78.

[Forina, 1991] M. Forina:

PARVUS- An Extendible Package for Data Exploration, Classification and Correlation

Institute of Pharmaceutical and Food Analysis and Technologies, Via Brigata Salerno, Genoa, Italy 1991.

[Fürnkranz, 1999] J. Fürnkranz:

Separate-and-Conquer Rule Learning

Artificial Intelligence Review **13**(1999 No 1)3-54.

[Galant, 1996] V. Galant:

GIMS - Decision Tree Learning System

W: Proc. 1st Polish Conference on Theory and Applications of AI, Łódź, 1996.

[Garner, 1995] S.R. Garner:

WEKA: The waikato environment for knowledge analysis

W: Proc. of the New Zealand Computer Science Research Students Conference, University of Waikato, Hamilton (New Zeland) 1995, s. 57-64.

[German, 1987] B. German:

Central Research Establishment Home Office Forensic Science Service Aldermaston, Reading, Berkshire RG7 4PN, 1987.

[Gerrits, Spreeuwenberg i Boekenoogen, 2000] R. Gerrits, S. Spreeuwenberg, M. Boekenoogen:

VALENS: A Knowledge Based Tool to Validate and Verify an Aion Knowledge Base

W: Proceedings 14th European Conference on Artificial Intelligence, 2000, s. 731 - 735

[Ginsberg, 1988] A. Ginsberg:

Knowledge-Base Reduction: A New Approach to Checking Knowledge Bases for Inconsistency and Redundancy

W: Proc. 7th National Conference on Artificial Intelligence (AAAI 88), St. Paul, MN, vol. 2, s. 585-589.

[Gonzales i Barr, 2000] A.J. Gonzales, V. Barr:

Validation and verification of intelligent systems

W: Journal of Experimental and Theoretical AI, Oct. 2000.

[Grzymała-Busse, 1992] J.W. Grzymała-Busse:

LEERS – A system for learning from examples based on rough sets

W: Słowiński R. (Ed.) Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, Dordrecht, Boston, London 1992, s. 3-18.

[Grzymała-Busse, 1994] J.W. Grzymała-Busse:

Managing uncertainty in machine learning from examples

W: Proc. 3rd Int. Symp. on Intelligent Systems, Wigry, IPI PAN Press, Warszawa 1994, s. 70-84.

[Grzymała-Busse, 1997] J.W. Grzymała-Busse:

A new version of the rule induction system LEERS

Fundamenta Informaticae **31**(1997)27-39.

[Grzymała-Busse, 2002] J.W. Grzymała-Busse:

MLEM2: A new algorithm for rule induction from imperfect data

Proc. 9th Int. Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU 2002, 1-5.07.2002, Annecy (France), s. 243-250.

[Grzymała-Busse, 2003] J.W. Grzymała-Busse:

A Comparison of Three Strategies to Rule Induction from Data with Numerical Attributes

W: Proc. Int. Workshop on Rough Sets in Knowledge Discovery (RSKD 2003), associated with the European Joint Conferences on Theory and Practice of Software 2003, Warsaw (Poland) 5-13.04.2003, s. 132-140.

[Grzymała-Busse, 2003a] J.W. Grzymała-Busse:

MLEM2—Discretization During Rule Induction

Proc. Int. Conference on Intelligent Information Processing and WEB Mining Systems (IIPWM '2003), Zakopane, 2–5.06.2003, Springer-Verlag, 2003, s. 499–508.

[Haendel, 2003] L. Haendel:

The PNC2 Cluster Algorithm. An integrated learning algorithm for rule induction

W: Clusterverfahren zur datenbasierten Generierung interpretierbarer Regeln unter Verwendung lokaler Entscheidungskriterien, PhD thesis, University of Dortmund, 2003.

[Harmelen, 1998] F. van Harmelen:

Applying rule-base anomalies to KADS inference structures

W: Decision Support Systems, Elsevier Science, vol. 21, no. 4, 1998, s. 271-280.

[Hippe, 1997] Z.S. Hippe:

Uczenie maszynowe - obiecującą strategią przetwarzania informacji w biznesie?

Informatyka **4** (1997)27-31, **5**(1997)29-33.

[Hippe i inni, 2003] Z.S. Hippe, S. Bajcar, P. Błajdo, J.P. Grzymała-Busse, J. W. Grzymała-Busse, M. Knap, W. Paja, M. Wrzesień:

Diagnosing Skin Melanoma: Current versus Future Directions

TASK Quarterly **7**(2003, No 2)289-293.

[Hippe, Grzymała-Busse i Mroczek, 2004] Z.S. Hippe, J.W. Grzymała-Busse, T. Mroczek:

Rules from Belief Networks: A Rough Set Approach

W: Tsumoto S., Słowiński R., Komorowski J., Grzymała-Busse J.W. (Eds.) Proc. 4th Int. Conference: Rough Sets and Current Trends in Computing, Springer-Verlag Berlin Heidelberg 2004, s. 483-487.

[Hippe, Knap i Paja, 2002] Z.S. Hippe, M. Knap, W. Paja:

Feasibility Studies of Quality of Knowledge Mined from Multiple Secondary Sources

W: M.A. Kłopotek, S. Wierzchoń, M. Michalewicz (Eds.); Intelligent Information Systems, Advances in Soft Computing, Springer-Verlag Berlin Heidelberg 2002, s.361-364.

[Hippe i inni, 2005a] Z. S. Hippe, J.W. Grzymała-Busse, P. Błajdo, M. Knap, T. Mroczek, W. Paja, M. Wrzesień:

Classification of Medical Images In the Domain of Melanoid Skin Lesions

W: M. Kurzyński, Puchała E., M. Woźniak, A. Żołnierek (Eds.), Computer Recognition Systems, Advances in Soft Computing, Springer-Verlag Berlin Heidelberg 2005, s. 519-526.

[Hippe i inni, 2005b] Z.S. Hippe, J.W. Grzymała-Busse, M. Knap, W. Paja:

Infoscience Technology: The Impact of Internet Accessible Melanoid Data on Health Issues

W: F.J. Smith (Ed.) Data Science Journal, Vol. 4, 20 September 2005, s. 77-81.

[Kohavi, Sommerfield i Dougherty, 1996] R. Kohavi, D. Sommerfield, J. Dougherty:
Data mining using MLC++: A machine learning library in C++

W: Tools with Artificial Intelligence, IEEE Computer Society Press, Los Alamitos, CA, (USA) 1996, s. 234-245 <http://www.sgi.com/Technology/mlc>.

[Kononenko i Bratko, 1991] I. Kononenko, I. Bratko:

Information-based evaluation criterion for classifier's performance

Machine Learning, **6**(1991)67-80.

[Krawiec i Stefanowski, 2003] K. Krawiec, J. Stefanowski:

Uczenie maszynowe i sieci neuronowe

Wydawnictwo Politechniki poznańskiej, Poznań 2003, s. 82-86.

[Kubat, Bratko i Michalski, 1998] M. Kubat, I. Bratko, R.S. Michalski:

Review of machine learning methods

W: Michalski R.S., Bratko I., Kubat M. (Eds.), *Machine Learning and Data Mining*, John Wiley & Sons, 1998, s. 3-70.

[Kulikowski, 1995] J.L. Kulikowski:

Podstawy reprezentacji wiedzy w systemach komputerowych

W: Traczyk W. (Red.), *Problemy sztucznej inteligencji*, Wiedza i Życie, **27**(1995)23-52, Warszawa 1995.

[Kulikowski, 2006] J.L. Kulikowski:

Model wnioskowania deontycznego w systemie doradczo-decyzyjnym z bazą wiedzy

W: Grzech A., *Inżynieria wiedzy i systemy ekspertowe*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2006, s. 163-172.

[Kwedlo i Krętowski, 2002] W. Kwedlo, M. Krętowski:

Learning Decision Rules using a Distributed Evolutionary Algorithm

TASK Quarterly, **6**(3)483-492, Gdańsk 2002.

[Laurent, 1992] J.P. Laurent:

Proposal for a validation terminology in KBS Validation

ECAI 92, John Wiley & Sons, Ltd., 1992.

[Ligęza, 1997] A. Ligęza:

Logical analysis of completeness of rule-based systems with dual resolution

W: Proc. of EUROVAV-97 - 4th European Symposium on the Validation and Verification of Knowledge Based Systems, Leuven, Belgium, 1997, s. 19-29.

[Ligęza, 2005] A. Ligęza:

Logical Foundations for Rule-Based Systems

Wydawnictwa AGH, Kraków 2005.

[Loh i Shih, 1997] W.Y. Loh, Y.S. Shih:

Split selection methods for classification trees

Statistica Sinica, **7**(1997 No. 4)815-840.

[Lunardhi i Passino, 1995] A.D. Lunardhi, K.M. Passino:

Verification of qualitative properties of rule-based expert systems

W: Applied Artificial Intelligence, **9**(1995)587-621.

[Michalski, 1969] R.S. Michalski:

On the quasi-minimal solution of the general covering problem

W: Proc. 5th Int. Symposium on Information Processing, Bled (Slovenia) 1969, s.125-128.

[Michalski, 1983] R.S. Michalski:

A theory and methodology of inductive learning

W: Michalski R.S., Carbonell J.G, Mitchell T.M. (Eds.), *Machine Learning: An Artificial Intelligence Approach*, Morgan Kaufman, 1983, s. 83-134.

- [Michalski, 1990] R.S. Michalski:
Pattern recognition as rule-guided inductive inference
 IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2, 4, 1990.
- [Michalski, Carbonell i Mitchell, 1986] R.S. Michalski, J.G. Carbonell, T.M. Mitchell:
Machine Learning: An Artificial Intelligence Approach
 Morgan Kaufmann, 1(1983), 2(1986).
- [Michalski i inni, 1986] R.S. Michalski, I. Mozetic, J. Hong, N. Lavrac:
The AQ15 Inductive Learning System: An Overview and Experiments
 Report No. UIUCDCS-R-86-1260, Department of Computer Science, University of Illinois, Urbana, IL, 1986.
- [Michalski i Kaufman, 1999] R.S. Michalski, K. A. Kaufman:
A Measure of Description Quality for Data Mining and its Implementation in the AQ18 Learning System
 Proc. of the Int. ICSC Symposium on Advanced in Intelligent Data Analysis (AIDA), Rochester 1999.
- [Minsky, 1974] M. Minsky:
A Framework for Representing Knowledge
 W: Winston P. (Ed.), The Psychology of Computer Vision, McGraw-Hill, 1975
- [Mitche, Spiegelhalter i Taylor, 1994] D. Mitche, C.J. Spiegelhalter, C. Taylor (Eds.):
Machine Learning, Neural and Statistical Classification
 Elis Horwood, London 1994.
- [Mitchell, 1997] T. Mitchell:
Machine Learning
 Mac-Graw Hill, Boston 1997.
- [Muggleton i De Raedt, 1994] S. Muggleton, L. De Raedt:
Inductive logic programming: Theory and methods
 Journal of Logic Programming, **19,20**(1994)629-679.
- [Murrell, Plant, 1995] S. Murrell, R.T. Plant:
A survey of tools for the validation and verification of knowledge-based systems: 1985-1995
 W: Decision Support Systems, Elsevier Science, Vol. 21, Number 4, December 1997, s. 307-323(17).
- [Nalepa, 2004] G.J. Nalepa:
Meta-Level Approach to Integrated Process of Design and Implementation of Rule-Based Systems
 Rozprawa doktorska, Akademia Górniczo-Hutnicza, Kraków 2004.

[Newman i inni, 1998] D.J. Newman, S. Hettich, C.L. Blake, C.J. Merz:
UCI Repository of machine learning databases
<http://www.ics.uci.edu/~mllearn/MLRepository.html>

[Nguyen i inni, 1985] T. Nguyen, T.J. Laffey, D. Pecora, W.A. Perkins:
Checking an Expert Systems Knowledge Base for Consistency and Completeness
W: Proc. 9th International Conference on Artificial Intelligence (IJCAI 85), Los Angeles 1985, 1(1985)375-837.

[Øhrn i inni, 1998] A. Øhrn, J. Komorowski, A. Skowron, P. Synak:
The Design and Implementation of a Knowledge Discovery Toolkit Based on Rough Sets – The Rosetta System
W: Polkowski L., Skowron A. (Eds.) *Rough Sets in Knowledge Discovery*, Physica Verlag, 1998.

[Paja, 2005] W. Paja:
RuleSEEKER – nowy system zarządzania wiedzą ujętą w formalizmie reguł składowych
W: R. Tadeusiewicz, A. Ligeza, M. Szymkat (Red.) *Computer Methods and Systems. Metody i systemy komputerowe w badaniach naukowych i projektowaniu inżynierskim*, Wyd. Oprogramowanie Naukowo-Techniczne, Kraków 2005, s. 367-370.

[Pappa i Freitas, 2004] G.L. Pappa, A.A. Freitas:
Towards a Genetic Programming Algorithm for Automatically Evolving Rule Induction Algorithms
W: Fürnkranz J. (Ed.), *Advances in Inductive Rule Learning*, Proc. 15th European Conference on Machine Learning (ECML) and 8th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD), Pisa (Italy) 2004, s. 93-108.

[Pawlak, 1982] Z. Pawlak:
Rough sets
W: *International Journal of Computer and Information Sciences*, 11(1982)341-356.

[Pawlak, 1991] Z. Pawlak:
Rough sets. Theoretical aspects of reasoning about data
Kluwer Academic Publishers, Dordrecht 1991.

[Pawlak, 1995] Z. Pawlak:
Wiedza i zbiory przybliżone
W: W. Traczyk (Ed.) *Problems in Artificial Intelligence*, Wiedza i Życie, Warszawa 1995, s. 9-21.

[Pindur, 2004] R. Pindur:
Agregacja wiedzy za pomocą reguł decyzyjnych o uogólnionej składni – metodyka i zastosowania

Rozprawa doktorska, Wydział Informatyki i Zarządzania Politechniki Poznańskiej, Poznań 2004.

[Preece, 1999] A. Preece:

COVERAGE: Verifying Multiple-Agent Knowledge-Based Systems

Knowledge-Based Systems, Elsevier Science, **12**(1999)37-44.

[Preece i Shinghal, 1994] A. Preece, R. Shinghal:

Foundation and Application of Knowledge Base Verification

W: D. E. O'Leary (Ed.), International Journal of Intelligent Systems, **9**(1994)683-701.

[Preece, 1989] A.D. Preece:

Verification of Rule-Based Expert Systems in Wide Domains

W: Shadbolt N. (Ed.), Research and Development in Expert Systems VI (Proc. Expert Systems 89), Cambridge University Press, 1989, s. 66-77.

[Quinlan, 1993] J.R. Quinlan:

C4.5: Programs for Empirical Learning

Morgan Kaufmann, San Francisco, CA, 1993.

[Quinlan, 1996] J.R. Quinlan:

Bagging, boosting, and C4.5

W: Proc. 13th National Conference on Artificial Intelligence, AAAI Press/MIT Press, Cambridge, MA 1996, s. 725-730.

[Quinlan, 1997] J.R. Quinlan:

Data Mining Tools See5 and C5.0

<http://www.rulequest.com/see5-info.html>

[Robertson i Sparck Jones, 1976] S. Robertson, K. Sparck Jones:

Relevance weighting of search terms

W: Journal of the American Society for Information Science, **27**(1976)129-146.

[Roiger i Geatz, 2003] R.J. Roiger, M.W. Geatz:

Data Mining: a tutorial-based primer

Pearson Education, Inc., San Francisco 2003, s. 34-36.

[Rousset, 1988] M.C. Rousset:

On the Consistency of Knowledge Bases: the COVADIS System

W: Computational Intelligence, **4**(2)1988, s. 166-170.

[Rutkowska, Piliński i Rutkowski, 1997] D. Rutkowska, M. Piliński, L. Rutkowski:

Sieci neuronowe, algorytmy genetyczne i systemy rozmyte

Warszawa, Łódź 1997.

[Salzberg, 1991] S. Salzberg:

A nearest hyperrectangle learning method

Machine Learning, **6**(1991)251-276.

[Setiono, 1996] R. Setiono:

Extracting Rules from Pruned Neural Networks for Breast Cancer Diagnosis
Artificial Intelligence in Medicine, **8**(1996)37-51.

[Słowiński i Stefanowski, 1994] R. Słowiński, J. Stefanowski:

Rough classification with valued closeness relation

W: Diday E. et al. (Eds.), *New Approaches in Classification and Data Analysis, Studies in Classification, Data Analysis and Knowledge Organisation*, Springer Verlag, 1994, s. 482-489.

[Soler-Toscano, Nepomuceno-Fernández i Aliseda-Llera, 2006] F. Soler-Toscano, Á. Nepomuceno-Fernández, A. Aliseda-Llera:

Model-Baded Abduction via Dual Resolution

W: L. Magnani (Ed.), *Logic Journal of IGPL*, Oxford University Press, **14**(2)2006, s. 305-319.

[Spreeuwenberg i Gerrits, 2002] S. Spreeuwenberg, R. Gerrits:

Requirements for successful verification in practice

W: *Proceedings Fifteenth Flairs conference*, 2002, s. 221.

[Stefanowski, 2001] J. Stefanowski:

Algorytmy indukcji reguł decyzyjnych w odkrywaniu wiedzy

Rozprawa habilitacyjna nr 361, Wydawnictwo Politechniki Poznańskiej, Poznań, 2001.

[Stefanowski, 1998] J. Stefanowski:

Rough Set based rule induction techniques for classification problems

W: *Proc. 6th European Congress on Intelligent Techniques and Soft Computing*, **1**(1998)109-113, Aachen 1998.

[Stolz i inni, 1993] W. Stolz, O. Braun-Falco, P. Bilek, M. Landthaler, A.B. Cogneta:

Color Atlas of Dermatology

Blackwell Science Inc., Cambridge (MA) 1993.

[Suwa, Scott i Shortliffe, 1982] M. Suwa, A.C. Scott, E.H. Shortliffe:

An Approach to Verifying Completeness and Consistency in a Rule-Based Expert System

AI Magazine, Vol. 3, Nr. 4, 1982.

[Śnieżyński, 2006] B. Śnieżyński:

Converting a Naive Bayes Model into a Set of Rules

Advances in Soft Computing **5**(2006)221-229, Springer-Verlag Berlin Heidelberg 2006.

[Towell i Shavlik, 1993] G. Towell, J. Shavlik:

Extracting refined rules from knowledge-based neural networks

Machine Learning, **13**(1993)71–101.

[Turmeaux i inni, 2003] T. Turmeaux, A. Salleb, C. Vrain, D. Cassard:

Learning Characteristic Rules Relying on Quantified Paths

W: Lavrač et al. (Eds.): PKDD 2003, LNAI 2838, Springer-Verlag, Berlin Heidelberg 2003, s. 471-482.

[Viktor, Engelbrecht i Cloete, 1998] H.L. Viktor, A.P. Engelbrecht, I. Cloete:

Incorporating rule extraction from artificial neural networks into a computational network.

W: Proc. of the Int. Conference on Neural Networks and Their Applications (NEURAP'98), March 11-13, Marseille 1998, France, s. 421-429.

[Viktor, 1999] H.L. Viktor:

The CILT multi-agent learning system

South African Computer Journal (SACJ), 24(1999)171-181.

[Viktor, Paquet i le Roux, 2003] H.L. Viktor, E. Paquet, G. le Roux:

Cooperative Learning and Virtual Reality-Based Visualization for Data Mining

W: Wang J., Data Mining: Opportunities and Challenges, Idea Group Publishing, London 2003, s. 55-79

[Weiss i Hirsh, 2000] G.M. Weiss, H. Hirsh:

A Quantitative Study of Small Disjuncts

W: Proc. 17th National Conference on Artificial Intelligence, Austin, Texas (USA) 2000.

[Winston, 1992] P.H. Winston:

Artificial Intelligence

Addison-Wesley Publ. Comp., Reading (MS, USA) 1992.

[Wróblewski, 2001] J. Wróblewski:

Adaptacyjne metody klasyfikacji obiektów

Rozprawa doktorska, Wydział Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego, Warszawa 2001.

[Zygmunt, 2001] A. Zygmunt:

Eksploracja danych metodą asocjacji (wybrane aspekty optymalizacji)

Rozprawa doktorska, Akademia Górniczo-Hutnicza, Kraków 2001.

[Żytkow, 2002] J.M. Żytkow:

Types and forms of knowledge: Rules

W: Klösgen W., Żytkow J.M., Handbook of Data Mining and Knowledge Discovery, Oxford Press 2002, s. 51-54.